DOCUMENT RESUME

ED 207 021 CS 006 268

AUTHOR	Davison, Alice, Ed.; And Others
TITLE	Text Readability: Proceedings of the March 1980 Conference. Technical Report No. 213.
INSTITUTION	Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY	National Inst. of Education (ED), Washington, D.C.
PUB DATE	Aug 81
CONTRACT	400-76-0116
NOTE	170p.
EDRS PRICE	MF01/PC07 Plus Postage.
DESCRIPTORS	Captions; Difficulty Level; *Discourse Analysis; Language Usage; *Measurement Techniques; Readability; *Readability Formulas; *Reading Research; *Research Problems; Television Viewing

ABSTRACT

The papers included in this collection represent as closely as possible the content and organization of the oral presentations delivered at a March 1980 conference on the use of readability formulas. The papers discuss the following topics: (1) an introduction to the conference, with general remarks on the uses and criticism of readability formulas; (2) the development of readability formulas and attendant problems in the validation of readability formulas; (3) readability formulas and the definition of the task of reading; (4) cases where readability formulas do not work well, including television captions for the deaf, elementary school tradebooks, reading comprehension tests, remedial reading texts, and basal readers; (5) lowering the reading difficulty level of texts intended for adults, with implications for "plain language" in legal documents; (6) readability formulas and the adaptations of texts; (7) the comprehension of captioned television; and (8) arguments against some uses of readability formulas, with suggested alternatives to readability formulas. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED207021

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

CENTER FOR THE STUDY OF READING

Technical Report No. 213

TEXT READABILITY: PROCEEDINGS OF THE MARCH 1980 CONFERENCE

Edited by

Alice Davison, Richard Lutz, and Ann Roalef

University of Illinois at Urbana-Champaign

August 1981

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02238

This publication of proceedings is supported in part by the National Institute of Education under Contract No. HEW-NIE-C-400-76-0116.

Ann Roalef Kantor is currently a graduate student in the Department of Linguistics, at Ohio State University, Columbus, Ohio.

CS006268

EDITORIAL BOARD

Peter Johnston, Chairperson

Roberta Ferrara
Scott Fertig
Nicholas Hastings
Asghar Iran-Nejad
Jill LaZansky
Jim Mosenthal
Ann Myers
Andee Rubin
William Tirre
Paul Wilson

Peter Winograd

Michael Nivens, Editorial Assistant

Table of Contents

Foreword--Alice Davison 1

1. Introduction to the Conference--Alice Davison 2

General Remarks 2

Uses for Readability Formulas 5

Earlier Criticisms of Formulas 6

Focus of the Conference 8

The Speakers 9

2. On the Validation of the Original Readability Formulas--Ramsay Selden 10

The Development of Readability Formulas 11

Some Widely Used Formulas 17

The Validation Problem 18

An Alternative Model 24

Discussion 26

3. Readability Formulas and the Definition of the Task of Reading--Margaret Griffin 31

Inadequacy of Reading Levels 31

Goals of Reading Instruction 33

Reader-as-subject vs. reader-as-object approaches 34

Cultural attitude and purpose 35

Model of Differences 36

Factors not considered by the formulas 36

Word familiarity and context 37

Syntactic Complexity 38

Leveling-Out 41

Zone-of-Proximal-Development Tutorial 42

Discussion 43

4. Jobs You Shouldn't Count on Readability Formulas to Do--
Bertram Bruce and Andee Rubin 45

Job 1: TV Captioning for the Deaf 48

Example 49

Problems 51

Assumptions 52

Job 2: Elementary-School Trade Books 53

Example 53

Problems 54

Assumptions 55

Job 3: Reading Comprehension Tests 55

Example 56

Problems 57

Assumptions 57

Job 4: Remedial Reading Texts 58

Example 1 59

Problems 60

Assumptions 61

Example 2 61

Problems 61

Assumptions 62

Job 5: Basal Readers 62

Example 63

Problems 63

Assumptions 65

Assumptions About Readability Measures and Their Use 65

Discussion 68

5. Lowering the Difficulty of Texts Intended for Adults: Implications for "Plain Language" in Legal Documents--Veda Charrow 74

Jury Instructions 74

The Document Design Project at AIR 75

The preliminary phase of the study 76

The task 76

Scoring the results 77

Interpreting the results 77

The Rewritten Instructions 80

Results 80

Applying the Flesch Formula to the two sets 81

Conclusion 81

Discussion 82

6. Readability Formulas and the Adaptations of Texts--Alice Davison and Robert N. Kantor 83

Adaptation and Sentence Length 84

Adaptations and sentence restructuring 85

Deletions of point of view and modality 89

Lexical Changes 93

Consequences of Formulas as Writing Guides 95

Reasoned adaptation 97

7. Comprehension of Captioned Television--Mark Seidenberg 98

Previous Research 100

Issues 101

Factors Affecting Caption Comprehension 105

Display conventions 105

Cues to sentence structure 106

Global structure cues 108

Lexical choice 110

Genres 112

Summary 115

Discussion 115

8 Final Discussion 118

Arguments Against Some Uses of Readability Formulas 118

Alternatives to Readability Formulas 123

Reference Notes 129

References 130

Footnotes 140

Tables 141

Figures 145

Foreword

The papers which make up this Technical Report are summaries of oral presentations on readability and readability formulas delivered on March 10 and 11, 1980 at the Center for the Study of Reading. There is some overlap in the material covered in different presentations, as our organization of the sections reflects the content of the particular paper presented by the authors named on it. Those who spoke are listed as authors. Acknowledgment is given to co-authors in the form of footnotes. Most of the papers were prepared from texts submitted by the authors, and all of them have been read and corrected by the authors. Several of the summaries of papers and all of the summaries of discussion were prepared from tape recordings of the procedures by graduate student members of the Text Analysis Group, Center for the Study of Reading. These included Jean Hannah, Margaret Laff, and Robert Salzillo.

The papers included here represent as closely as possible the content and organization of the oral presentations, in a more readable format than a verbatim transcript. We would like to emphasize that the purpose of the conference was to raise a number of issues for discussion and to present a spectrum of ideas and viewpoints from which readability formulas could be judged or criticized. We have not tried to make the papers exhaustive summaries of all that has been done on a certain subject or

to represent only the most "correct" and orthodox positions on any topic. We feel that the views expressed here, while diverse and in some cases programmatic, will be useful in provoking discussion and reexamining assumptions about readability formulas, perhaps in defining research which might lead to a better understanding of what makes things difficult to read.

I. Introduction to the Conference--Alice Davison

General Remarks

In this introduction, I want to describe how we came to consider the questions raised and discussed during the conference. Basically our interest grew out of research which was originally not directed toward the criticism of readability formulas, a study done at the Center for the Study of Reading by the linguists in the Text Analysis Group. In this study, we set out to analyze the structures of texts used as reading materials for children. In doing this, we hit upon the idea of comparing the version of a text used for this purpose with the original text from which it had been adapted, which might involve a change in structure and also some simplification to make the adapted text easier to read. Somewhat to our surprise, we found that the changes made in adaptation were of very great interest, and our study was focused on characterizing the kinds of changes made and the possible motivation behind them, assuming that any changes made were intended to make the text easier to read.

The results are discussed in Section 6 by Robert Kantor and Alice Davison, but here I want to note one of our conclusions, based on the following points. Words in the text which were not necessarily obscure in context, or even very infrequent, were changed in adaptation to shorter words or more familiar words. Changes made in sentence structure, such as splitting a long sentence into several sentences, tended to make sentences shorter throughout the text and in the overall average sentence length. Such changes were made even when the original sentence structure did not seem to us to present any particular difficulties. The commonly used readability formulas measure vocabulary frequency or number of syllables per word, plus sentence length--exactly the features of the texts which were modified without apparent motivation within the text. It seems that the changes were made in order to make the adapted texts "meet" a certain level of reading skill defined by these formulas.

Other kinds of changes were made in order to clarify or simplify the texts. From these, we were able to isolate some other features of texts, less easily measurable by formulas, which in our view contributed to text comprehension as much as, or more than, sentence length and vocabulary complexity (see Section 6, Davison and Kantor). In fact, some of the cases where vocabulary was simplified or sentences were split up tended to make the text less easy to comprehend, because the relations between the parts of the text were less clearly expressed.

One of the main conclusions of our text analysis study was that readability formulas, whatever their value as predictive, statistically based measures, are not particularly helpful in directing writing or adaptation. They fail to address the central issue: What is it about the language of a text which might normally make it easy or hard for a reader to understand the text? (See von Glaserfeld, 1970-71, for a lucid discussion of what might be meant by syntactic complexity and how it affects comprehension.)

There are increasing numbers of cases where the answers to this question are necessary. If the language of some texts is such that people who need to be able to understand the texts are, in some large numbers, unable to do so, then it would be useful to have some set of principles for changing the language of the text to make it comprehensible to people reading at some specific level of reading achievement. Here a systematic ambiguity in the term readability becomes apparent. Sometimes the term is used to refer to text types or styles of writing that people find attractive and hence easier to read than other kinds of writing (e.g., Flesch, 1948). But nowadays the main concern is with what people are able to read, under normal circumstances.

The problems of major concern today have to do mainly with texts that people must read and be able to understand, such as school and college-level textbooks, tax forms and instructions, legal documents and similar texts on technical subjects whose contents must be clearly expressed in language understandable to the general public. Of course, not all of these areas of concern are directly related to one another.

Some important differences having to do with tasks and with child versus adult readers are discussed in the following sections. Nevertheless, these are all areas where writers and publishers have looked to readability formulas for some sort of guidance.

Uses for Readability Formulas

Formulas were originally devised as predictive averages which would be useful in ranking in order of difficulty a number of books for use in a particular school grade. They were to be simple enough to be applied by people without extensive training, and to be applicable to texts of many different types. Thus, as the creators of formulas have always pointed out, the formulas were intended for certain specific uses, for example, estimating the difficulty of a text in relation to some other text. It seems clear that formulas, which measure objectively definable features such as sentence length and more or less definable features like vocabulary complexity, may reflect but do not characterize the real factors which contribute to ease or difficulty in reading a text. Formulas therefore cannot serve as guides for writing, nor were they meant to. Proponents of formulas such as Klare (1963, 1974-75) point out that the use of formulas presupposes the existence of a well-formed text to begin with, one which has been put together by a writer using whatever common sense and writing skill he or she possesses, including the ability to present a subject coherently, etc. The valid use of formulas thus presupposes a number of

properties of text besides sentence length and vocabulary complexity, and these properties are undefined and unexamined by the authors of formulas.

Although formulas were not designed to be used as guides for writing, they are used in this way, though nobody really knows how widely. Even if not actually used as guidelines, they have the indirect effect of guidelines if they are used to define a standard reading level which a given text must meet. Publishers and writers who must create reading materials at levels of reading difficulty defined by formulas are under great pressure to make their materials conform by altering vocabulary choices and sentence length.

Earlier Criticism of Formulas

At the end of the adaptation study, which compared two versions of the same text, we felt a glow of honest pride about what we had discovered regarding changes in adaptations and about our conclusions, which were critical of the use of readability formulas. Then we discovered that for at least the past ten years, a number of other people have come to the very same conclusions about formulas (Botel & Granowsky, 1972; Dawkins, 1975; Endicott, 1973; Gourley & Catlin, 1978; Nelson, 1978; Reddin, 1970; Selden, 1977; von Glaserfeld, 1970-71; Fagan, Note 1; Schmidt, Note 2), and have proposed various alternative measurements, mostly in very programmatic form. The common denominator of these proposals is some method of measuring syntactic complexity more directly than just

through sentence length. Such measures at least make the measurement of complexity more text specific, although they do not attempt to characterize such larger notions as presentation of ideas, definition of topic, etc.

We were surprised that the articles which preceded our work had not had very much effect on the way that readability formulas are perceived. Formulas continue to be in widespread use, both as measures and guidelines, in spite of awareness on the part of some researchers of their very serious drawbacks as reliable instruments. One probable reason for this is that formulas have very strong economic advantages in that they can be used without special equipment by people without special training. The more direct measures of syntactic complexity and other features require some specialized knowledge of language structures on the part of the analyst and probably take more time to apply because the text must be very closely examined and analyzed. Readability formulas also have the advantage of being well known, familiar, and already part of the scheme of things, while the alternatives are speculative and relatively untried.

Nevertheless, the disadvantages of using traditional readability formulas, particularly for revising or writing texts, seem obvious and compelling to anyone who seriously considers them. Much energy has been invested in refining and simplifying formulas, with fewer or more easily defined variables (see surveys such as Klare, 1963, 1974-75). Not as much attention seems to have been paid to the much more

central and interesting questions of what makes a text difficult to process and comprehend, and exactly how the language in a text contributes to difficulty.

Focus of the Conference

If the answers to these questions were found, there might have been a number of political and economic consequences. But leaving aside such considerations, we approach formulas from two critical points of view in this conference, in order to establish what the intellectual issues are. First, to what degree are readability formulas really adequate and justified as devices for measurement--what is their conceptual basis, how were they initially validated, to what uses can they legitimately be put? Where they fail to be adequate or appropriate, what are valid and compelling arguments against them?

Second, what alternatives are there for readability formulas, especially for particular purposes, such as writing and revising texts? Some proposals have included taxonomies of structure types (Botel & Granowsky, 1972; Dawkins, 1975), investigations of reading as a left to right parsing of sentences or longer structures (Richek, 1976; von Glaserfeld, 1970-71; Fagan, Note 1), weighting of different general structures (Endicott, 1973; Reddin, 1970; Schmidt, Note 2; Selden, Note 2), measurement of difficulty relative to context in the discourse (Gourlay, 1978; Gourlay & Catlin, 1978), and so on. What would constitute compelling evidence for such descriptions as improvements on

readability formulas? How could these alternatives be shown to be better guides to writing than formulas are?

Please note that the presentations at this conference are critical of readability formulas. This orientation is intentional; we feel that readability formulas have had a wide hearing throughout their history and that they do not need defenders here simply for the sake of balance. But our purpose is not to condemn readability formulas out of hand. Rather, it is to clarify the issues surrounding readability formulas and to define the logic behind various views. Often it seems that defenders of formulas and critics of formulas are really talking about separate issues, for example, statistically valid predictive averages versus explicit, internally consistent models of language processing. It is important to place both criticisms and defenses within the right contexts of discourse and not to generalize from one domain of discourse to another. The outcome of discussion can have no value if the issues are not debated within clearly defined contexts.

The Speakers

The following people presented papers during the two-day conference:

Bertram C. Bruce	Bolt, Baranek & Newman, Inc.
Veda Charrow	American Institute for Research
Alice Davison	Center for the Study of Reading (CSR)

Margaret Griffin	University of California, San Diego
Robert N. Kantor	Now at Dept. of English, The Ohio State Univ.
Andee Rubin	Bolt, Baranek & Newman, Inc.
Mark Seidenberg	Now at McGill University
Ramsay Selden	National Institute of Education

The following participants led discussions of presentations:

Rob Tierney	CSR
Georgia Green	CSR & Dept. of Linguistics, Univ. of Illinois
Jean Osborn	CSR

All of the people who spoke at the conference have had to confront readability formulas in some area where it is important to make a text comprehensible to a particular audience of readers. With the points that they raised, and the discussion of them, we hope to make some useful definitions of the issues and arguments bearing on them.

2. On the Validation of the Original Readability Formulas

Ramsay Selden[1]

This talk focuses on: (a) a brief history of the development of readability formulas, with a critical look at how researchers settled on the most widely used variables; (b) a variety of problems and circularities existing in the validation of readability formulas; and (c) Selden's own research on readability, which represents a departure from traditional use of readability formulas and which serves to point out what is wrong with readability formulas and how they are used.

Basically, readability formulas involve empirical counts of three or four characteristics of a text, which are multiplied by weighted coefficients, with the resultant numbers being summed to give an index of the readability of a piece of text. A readability index number for a given text is typically based on three samples of 100 words each from that text. Readability formulas are intended to be indices of the difficulty of a passage, but they were never intended to provide specifications of text characteristics that contribute to text difficulty.

The Development of Readability Formulas

Readability research probably began in earnest with the publication in 1920 of E. L. Thorndike's Teacher's Word Book. This was a listing of 10,000 vocabulary items stratified by their frequency of appearance per million words in several corpora of language data. It represents a systematic effort to provide a tool for the estimation of the difficulty of words which can assist in the teacher's introduction of appropriate material to students learning to read. Systematic analysis of text was not new, of course; for instance, Klare (1963) cites studies in the 1800s aimed at classifying text by its era of publication according to the mean length of its sentences. Prior to that, Herbert Spenser and language philosophers had speculated on the characteristics of text that entered into readability and on the concept of readability itself. Still, Thorndike's wordlist reflects a

perception of the pedagogical usefulness of examining reading materials and validly estimating the difficulty with which they may be read. Lively and Pressey (1923) published a study of vocabulary burden based on the Thorndike wordlist and found the mean frequency strata of words in text to be associated with the grade level in which the textual materials were to be used. This was an important precedent, as that has literally been the paradigm for readability technology. Using his own frequency-stratified wordlist, Dolch (1928) found that books within a single grade varied widely in estimated reading difficulty.

Around 1930, a substantial amount of experimentation was begun concerning the variables used as predictors of readability. In 1928, Vogel and Washburne predicted the tested reading ability of subjects who liked a book on the basis of its proportion of unique words to total words (type/token ratio), proportion of prepositions, proportion of words not on the Thorndike list, and the proportion of simple sentences. Most useful were the Thorndike measure of vocabulary difficulty and the proportion of simple sentences. Lewerenz, in six studies from 1929 to 1939, found the difficulty of text associated with the initial letter of words in the text ("I" and "E" were hard, "W," "H," and "B" were easy), with the presence of words of Greco-Roman derivation, with high-frequency vocabulary, with the presence of sensory words, and with the presence of polysyllabic words.

Generally, attention in investigations from 1930 to the present has split into considerations of word characteristics and considerations of sentence characteristics as predictors of readability. There has been virtually no work on the characteristics of the message structure of texts until the work of Meyer (1975) and Kintsch and Vipond (1979) in the last few years. Summary works such as the one done by Klare (1963) indicate a set of word features which have consistently been shown to be useful in estimating readability: vocabulary familiarity, based on a frequency-stratified word list developed from samples of text, prose, letters, or other bodies of written language; word length measured both in letters and in syllables; and word spelling, specifically the presence of certain graphemes in certain positions in words. As one notable exception, Rudolf Flesch has not included vocabulary familiarity in his readability formulas; instead, his 1948 work details two subscales, a "reading ease" scale based on word length and sentence length and a "human interest" scale based on the number of personal references in the text. His later studies described developments of the readability measures based on the proportion of affixed and abstract words (Flesch, 1951, 1954). Flesch's formulas are based on the premise that words in the text which appeal directly to the reader's attention and which provide content which is concrete as opposed to abstract will be more interesting and more easily read. However, as Klare points out (1963, pp. 19-20), this premise probably

holds only for certain types of text, such as fiction. Across all types of text, including technical, narrative, and expository writing, the words hypothesized to be readable by Flesch would often be unanticipated and could be expected to disrupt reading, making it more difficult. Significantly, the Flesch "reading ease" formula, based only on sentence length and word length, was cited by Klare as the single most widely used formula.

Researchers have experimented with a number of sentence characteristics as determinants of readability: numbers of prepositional phrases, degree of embedding or subordination, presence of structures such as passive verb forms. In most studies these variables have shown relatively weak correlations with criteria of readability, beyond the correlation provided by sentence length alone. Since these sentence-structure variables are highly correlated with sentence length (highly subordinated variables tend to be long), they have not endured as predictors of readability.

To summarize, since its beginning in the early 1920s, readability research has investigated a variety of textual features, but has most consistently preferred a small set, which has displayed enduring usefulness. Investigation also explored alternative methods for measuring criterion measures, predictor measures, and assigning weights to predictors in the formulas. By the mid-1950s, most formulas were working with just three variables--vocabulary familiarity, word length, and

sentence length. In most widely used formulas, a combination of two of these three variables is used.

There seem to be two primary reasons why readability formulas settled on these three measures. First, formulas containing just two of the variables characteristically show correlations of .7 and .8 with criterion readability scores (Dale & Chall, 1948; Fry, 1968; Spache, 1953). This indicates a good deal of predictive power, from 50% to 65% of the variance in criterion readability, and there is little increase in predictive power to be gained through the incorporation of additional variables. This is pointed out in Selden (1977); since readability formulas fail to account for syntax except for what may be incidentally measured by sentence length, this study attempted to add syntactic variables to the formulas. The research tested whether or not pieces of texts containing sentences that use relatively common surface strings would be more readable than texts with sentences containing relatively uncommon or infrequent surface strings. When a measure of frequency of surface strings was added as a variable to another formula, it was found that this added significantly to the predictive powers of the formula for older readers dealing with more difficult texts, but did not have much effect for younger readers. While it was valid to assume that the variable was significant to some degree, the question was, Under what conditions is it important to account for it? The addition of a variable has to be weighed against the payoff.

The second reason why readability formulas settled on those three basic measures is that it has been desirable for readability formulas to be convenient to use, requiring counts of textual variables which can be compiled by a classroom teacher or publishing house staff without a great deal of time and complicated tabulation. This was probably more true in the past, before there were computerized text-analysis capabilities. Also, the desire for convenience depends on the premise that classroom teachers use readability formulas on some widespread basis to evaluate texts and to match them up with students. But, in the only known study of the use of such formulas in the classroom, part of a survey done by Jean Chall and Edgar Dale in the mid-1950s, very little use of readability formulas by teachers was found. Indeed, it seems logically and operationally unreasonable for teachers to use readability formulas. Good reading specialists and reading teachers tend to work with two vast domains of intuitive nonempirical information in matching children and texts. One domain is knowledge about the child, which goes far beyond standardized reading test scores to include information on what kind of reading materials the child has read in the past, what his or her interest might be, more detailed indications of what his or her reading problems might be, and so on. A second domain includes the vast amount of information about reading materials which goes beyond what readability formulas tend to measure. Reading specialists and reading tutors have a

repertoire of books they use with children, and for which they know the approximate difficulty level, what the story is about, basically which words in the book a particular child might need special help with, and so on. All of that information is infinitely more useful in matching the book with the child than a technological assessment of the material compared with the child's reading test score.

Some Widely Used Formulas

Dale and Chall (1948), Dolch (1948), Fry (1968), Gunning (1952), Spache (1953), and Sticht (1972, 1975) have devised formulas based on the three basic measures mentioned above (vocabulary lists, word length, sentence length, or a combination of two of these measures), and which have both properties of being easy to use and of bearing a high degree of indicated predictive power.

Dale and Chall

Comprehension = .1579 (% words not on Dale-Chall list of 3000 common words) + .0496 (words/sentence) + 3.6365.

Gunning

Readability Index = .4 (mean sentence length + % words over two syllables).

Spache

Publisher's grade level = .141 (mean sentence length) + .086 (% words not on Dale list of 769 easy words) + .839.

Dolch

Publisher's grade level = average of table values for median sentence length, 90th percentile sentence length, and % words not on Dolch list of 1000 easy words.

Fry

Publisher's grade level = intersection of values for sentence length and word length measured in syllables on the Fry Readability Graph.

FORCAST (Sticht, 1975)

$$\text{Comprehension} = 20 - \frac{\text{number of one-syllable words.}}{10}$$

The Validation Problem

The problem of validating these formulas needs to be examined in closer detail. Most formulas are validated against one of two criterion measures: publishers' assigned grade level for passages from reading texts, or the norms for the McCall-Crabbs comprehension passages. The Dale-Chall formula was validated against the McCall-Crabbs (1925a, 1925b, 1961) comprehension norms; Gunning, against other readability formulas; Spache, Dolch, and Fry against publishers' assigned grade level; and FORCAST (Sticht, 1975) against comprehension questions. Formulas are developed by conducting multiple regression analyses on a set of passages (e.g., 50) using a range of predictor variables and working out the best set of weights and constants to estimate the criterion.

The publishers' assigned grade level simply reflects where a passage occurs in a basal series (usually 1.1 through 6.2). In the

past, passages were assigned to their various levels by common sense, writer feedback, teacher feedback, and other information, with formulas being used as a check. Since most basal series go only through level 6.2, the formulas are then based on elementary passages, even though they are supposed to be predictive up to adult reading levels. Besides this problem, there are circularities that plague the validation of formulas against publishers' assigned grade level. First is a pedagogical circularity in that, to some extent, the skills that are emphasized in instruction of school children correspond with the kinds of text characteristics the formulas are geared to measure. The Harris-Jacobson (1974) formula probably goes the farthest in doing this. In its four-variable version, it measures sentence length, word length, word familiarity based on frequency of use in basal reader programs, and word "decodability" based on phoneme-grapheme correspondences. Word frequency and decodability are both heavily taught in basal reader programs. Over time, we approach a technology that on the one hand teaches certain skills and on the other hand evaluates the demand for those skills in texts as indicators of the readability of the text. There is a danger that after a while what is measured is simply what is taught, bearing a partial relation to real-world literacy.

There is a more superficial level of circularity involved in using publishers' assigned grade level as a criterion. Simply, publishers use

readability formulas to some extent to control at least the grade placement of passages, if not the revision or the initial writing of the passages themselves. Thus, formulas can become self-fulfilling prophecies in continuing to be applied to passages that are written to conform to them.

The other criterion against which formulas are validated is an empirical measure of the comprehensibility of passages. The most prominent of these are the McCall-Crabbs (1925a, 1925b, 1961) comprehension passages. These are short paragraphs, mostly fiction or expository nonfiction, followed by 12 to 15 multiple-choice content questions. Because of the relationship between the questions and the passages, the passages are felt to represent a set of materials for which there is some empirical indication of what is learned from them. These passages were administered to a large number of children of different grade levels. Two numbers were produced: the C50 criterion of readability, which gives the average grade level of children who got at least 50% of the questions correct for a particular passage, and C70, the average grade level of those who got 70% correct. Advantages of the McCall-Crabbs passages are that they were normed on a large population, and there seems to be some empirical basis for grade placement of the passages. There are several problems with this test, however. Are the questions more difficult than the text itself? Are the questions an accurate measure of learning from the text? Do the

questions measure something a child has learned from the text or has brought to the text? Are there problems with testing error, fatigue, and item error? Further, the original norming did not involve a wide range of readers, text types, or tasks, and there is still a pedagogical circularity in that the passages are similar in many respects to those from basal readers. And finally, problems arise in validating a modern formula against highly dated materials and norms. In one attempt to update the norms, however, using contemporary children reading the same passages used in 1925 and answering the same questions, no significant differences were found except at the 12th-grade level (Jacobson, Kirkland, & Selden, 1978).

Cloze tests have been used as independent measures of readability, too. The cloze procedure involves deleting every fifth word from a passage, with the subjects' accuracy in supplying the exact missing word indicating the readability. One of the objections to it is that synonyms are not counted as correct, although it is claimed that this does not affect its validity. Another objection is that this test does not involve the same psychological processes as are involved in reading a text and performing some task afterwards; i.e., the cloze task involves looking back in the text for clues to the deletions so that it is more like solving a verbal puzzle than like reading a text in a linear fashion. The correlations may be high because the verbal problem-solving skills that are useful in doing a cloze procedure on a passage may be

held by readers who are proficient in various other kinds of comprehension tasks.

Researchers at NIE conducted a study that serves to illuminate what is wrong with readability formulas and how they are used. The study involved the analysis of the difficulty of the federal income tax package in terms of the speculated ability of the average adult to read the package and use the information in it. The General Accounting Office insisted on some kind of readability measure being used to verify that (a) the original package was too difficult and (b) that, with the recommended changes, the resultant tax package met certain levels. The estimated difficulty was lowered according to the Dale-Chall, Harris-Jacobson, Fry, and FORCAST formulas. There was a good deal of variability among the formulas about the estimated level, as many as four grade levels difference, with the average grade level being about 12.5. Other kinds of analyses, such as detailed linguistic and visual format analysis, were carried out to point out difficulties in the text without reliance on the formulas. For example, there was extensive use of left-branching structures in the forms, such as preposed <u>if</u> clauses, which psycho-linguistic research (e.g., Palmer, 1974) has shown to be difficult to process, and which are not accounted for in any formulas, except to the extent that sentence length reflects this. If these structures were moved to the other end of the sentence, the sentence might be easier to process even though it would be the same length.

It was estimated that one-third of the adult population confronted with the original tax forms would have difficulty reading them, based on what little information there is on how well adults read. A set of revisions was proposed, moving conditional clauses to the end of sentences, replacing some of the technical vocabulary, and doing a good deal of work with the visual format. These revisions lowered the readability level by three to four levels (to about 8.5), mainly by changing factors that are not directly measured by readability formulas. At the same time, the IRS independently revised the tax materials, lowering the readability estimates by four grade levels (to 8) mainly by reducing sentence length and word difficulty, variables which are measured directly by the formulas. Nevertheless, when a reporter gave these IRS-revised forms to some bright eighth graders in San Jose, very few of them could fill out the form correctly, given simulated data. (Unfortunately, there was no controlled field testing of the revised NIE tax forms on a sample group of potential users.)

The failure of eighth-graders to adequately comprehend text evaluated at eighth grade level points out what is wrong with using readability formulas for evaluating and controlling the difficulty of a variety of types of text. One problem is the publisher's use of the formulas, not as indicators of the general difficulty of a text, but as indicators of what changes should be made to lower the difficulty of particular texts. This not only does not work, but it also ignores

the explicit warnings of the formulas' inventors that they are not valid for such purposes. A large problem with the IRS's attempt is that the formulas used are based on the abilities of elementary school children and the properties of elementary school texts, while the tax materials were written to correspond to an adult level (8-12). There are striking differences between the range of adult readers' abilities and that of learning readers in elementary school. There are also radical differences between the type of text used in tax instructions and in narrative and expository prose, and there are striking differences between the tasks of reading instructions and following them on a form, and of reading a paragraph and answering questions about it. Readability formulas are not sensitive to differences in type of text, type of reader, or type of reading task.

An Alternative Model

This brings us to an alternative model for the problem of measuring readability developed by myself and Janice Reddish of the American Institute for Research. The objectives behind developing this model are (a) to provide some basis for estimating and evaluating the readability of a piece of text, and (b) to come up with a technology that is self-enriching in order to increase our understanding of what makes things readable. This new model is not a formula, but a more organic approach to the problem. It involves first identifying what the function or intended purpose of the particular piece of text is: i.e., what task

is involved, who is to perform it, and what genre the text is. Those three dimensions seem to be useful and necessary ways to classify the function of a piece of text. It is necessary to have some sense of what kind of reader the piece is being presented to, since readers differ. Also, the task is very important, since different tasks (getting the gist of something, memorizing details, following instructions, etc.) may involve different types of reading skills. Third, materials of different genre differ in their structure and applicability and can be used by different people to do different things (see Figure 1).

Insert Figure 1 about here.

If one starts out, then, by specifying the function of the text, there is a body of scientific and practical literature one can draw on to determine what characteristics the texts should have in order to meet their functions. In the scientific literature, there are 50 years of readability research and 100 years of psychological research on how characteristics of texts interact with characteristics of people. There is a practical body of literature in the fields of technical writing, graphic design, and educational technology. To some extent, the design of a piece of text can be controlled experimentally; i.e., one can identify, control, and keep track of what is being manipulated in the document. The next step is to test the new piece of text through

a field situation that tests the adequacy of the text against its original function. That evaluation is used to verify whether or not the design precepts hold.

This system does two things. First, it gives a basis for pragmatically and realistically telling how difficult a document is for its intended purpose, because an operational field test has been run. Second, this cycle of drawing on the research literature and the practical guidelines, systematically controlling the design of texts, and field testing the texts has implications for enriching the validity of the scientific and practical guidelines. This is one alternative to readability formulas. It does not account in detail for the way a given document is going to interact with the psychological and personal characteristics of an individual user. What it does do, is provide a technology for systematically designing and controlling documents.

Discussion

Green: The AIR alternative model to readability formulas just described makes a great deal of sense, and in fact, fits in with some research which I would like to conduct. My idea was to conduct field tests of trade picture books which children in kindergarten or first grade like to read, but which are rated higher (about third grade level) by standard readability formulas, to determine if such trade books really are that difficult. A number of trade publishers have "Easy

Reader" series which are advertised as vocabulary-controlled and sometimes controlled by readability formulas as well. Typically, such trade publications are much more interesting than basal readers.

Another point brought up by Selden was that proponents of readability formulas suggest that they be used by teachers and assume that they will be used by teachers to fit materials to their students. While I concur that teachers probably have not got the time--if in fact they have the inclination--to conduct such work, I am opposed to such a suggestion on principle. The philosophy behind that is to discourage teachers from relying on their instincts, but rather, to rely on the judgments of "scientists," which are labelled as valid and objective. This philosophy explains, at least in part, why teachers' manuals are often three times the length in pages and many more times the length in text of the text being taught.

Also brought out by Selden's presentation was the comment that applying Flesch's (1948) reading case/human interest scale to technical documents might result in bizarre texts. If the Art of Readable Writing (Flesch, 1949) is any indication, this is indeed the case, since Flesch practices what he preaches. The text is very much like spoken language, and is filled with direct commands and colloquialisms.

Selden: One way in which a lower readability score can be reached is by avoiding semi-colons and replacing them with periods--which would

appear to be one of Flesch's strategies. Note also that replacing technical vocabulary with simple words may forego accuracy, and may not account for the fact that the audience for which a text is written may be familiar with the technical words. This would be the case within the Government, for example.

Seidenberg: Several issues have been brought up that are hardly debatable. One is that readability formulas have not been validated in a valid way. A second is that people can abuse readability formulas in various ways. For instance, they might use them as a guide to writing texts, instead of as a way of getting some additional information about something which has already been written. However, a lot of issues remain open to question. For example, improper validation of the formulas does not necessarily invalidate the entire enterprise of readability formulas. There may still be a reason to want formulas which will be easy to apply to texts, and there might still be a theoretical basis on which to develop valid formulas. The AIR type of alternative is acceptable only if there is a large amount of time and money available for the field testing and evaluations. Certainly, such testing must be done for at least some texts some of the time. I would expect publishers to react instead by suggesting that the validation tests be done properly on already existing formulas or that new formulas be developed to perform the same functions as those already existing formulas.

Selden: In my dissertation research (Selden, 1977), I point out that the formulas fail to take into account the syntax of the sentences at all. Only sentence length has prevailed as a variable which indicates syntactic difficulty. Others had noted in the 1930's and 1940's that such features as number of prepositions, number of independent clauses, as well as other grammatical features seemed to be associated with readability. In the 1970's, a great deal of Chomskyan research was done to investigate the effect of the number of transformations implicit in the surface string on readability. However, nobody had investigated how frequency of occurrence relates to how difficult the syntax is for the reader. My hypothesis was that texts which contain relatively common surface strings would be more readable than texts which contain relatively uncommon surface strings. I went through an elaborate procedure to measure this effect and to construct a readability formula variable that would account for this. What I found was that addition of such a variable (frequency of occurrence of surface structures) added significantly to the predictive power of the formula. The effect was much greater for older readers than for the younger readers. I needed a curvilinear transformation of the variable in order to get a significant effect, even for the older readers. The marginal change was extremely low. I was left with the pragmatic conclusion that while this feature may be a valid part of readability, it might not be important to account for it under all

conditions. Certainly writing skill and talent and plain good writing would be at least as important.

Adding variables to formulas is an extremely bulky process. With the advent of computerized text analysis, the addition of new variables is becoming more feasible.

The real puzzle concerns the validation aspect of the problem. Why haven't additional validation studies been performed on these formulas? How much is involved in taking a different set of passages, applying the formulas, and comparing this to other criteria which indicate the difficulty of the passages? Virtually none of this kind of research has been conducted.

Seidenberg: Given the serious questions about the validation procedure, it is difficult to evaluate the meaning of studies which show the effect of additional variables to be minimal. On psycholinguistic grounds, a variable such as frequency of occurrence might be expected to make considerable difference.

Selden: This is in part a problem of regression equations. In my study, I would have preferred not to have controlled sentence length. One might then wonder whether or not sentence length was interacting with frequency of occurrence. Therefore, it was necessary to control sentence length. One effect washes out the other. The answer to such dilemmas is experimental studies. Building passages which are controlled in all respects save one enable the researcher to see how comprehension differs when manipulating the variable.

3. Readability Formulas and the Definition of the Task of Reading

Margaret Griffin[2]

In working with children, it becomes clear that the levels defined by readability formulas are misconceived in that they are based on misconceptions about the task of reading. The formulas are least useful at the lowest and highest levels: At levels near the top, the formulas characterize as readable what we would generally call poor writing, and at the lowest level (Grade 1 and under) the Spache formula (1953) is the only formula that can be used. To define reading as progression through reading levels is inadequate.

In spite of these facts, readability research has an impact on instruction in several ways. First, teachers may accept the suggestion that in the primary grades the goal of reading instruction should be to get children to read longer sentences with bigger words--a strange goal for reading. Secondly, the formulas have an effect on the materials produced for reading instruction, since they are widely used by publishers and state textbook committees.

Inadequacy of Reading Levels

The inadequacy of defining reading as progression through readability levels can be illustrated by looking at a story composed by a kindergarten child, as reported in Vivian Gussin Paley's Wally's Stories (1981):

A little boy lived all alone in a deep forest. When he wanted to know a word, he asked lions and tigers and wolves. They told him pretend words, because he couldn't speak animal language. One day he saw a lady and a man who didn't have a little boy. "What language do you speak?"

"Animal pretend talk."

"That's OK, 'cause we can teach you people's language. Which one do you want to learn?"

"English."

"Good, because that's our language. What words do you want to know?"

"Lion, tiger, and wolf."

"You already know them. You just said them."

"Then animal pretend talk must be English."

So they lived happily ever after. But the man and lady knew some words the boy didn't know, so they did have a lot to teach him.

The following is an example of a typical story to be found in a basal reader for beginning readers:

Bill said, "Look here, Ben.
Look at the ducks."
"Here, ducks," said Ben.
"Look at this.
Get it, ducks."

(Ginn 720 Level 3; see Clymer, Parr, Gates & Robison, 1976, pp. 11-12)

When Wally's Story is compared with the kinds of stories presented to beginners in basal readers, it becomes clear that the notions of ease defined by readability formulas do not correspond well to children's

actual abilities. The child writer exhibited more awareness of connectives, for example. In fact, the texts written for children may actually confuse them as to what reading is for.

Goals of Reading Instruction

This brings up the major issue of the goals of reading instruction. Walmsley (1981) argues that the goals of reading instruction are ideological by nature. There are three basic ideological positions on the goals of reading instruction: cultural reproduction, utilitarian, and interactive. Those who see the goals of reading in terms of cultural reproduction are concerned with the eventual ability of the reader to read the works of the culture that are deemed important--Faulkner or Shakespeare, for example. The second viewpoint, the utilitarian orientation, is directed toward the ability of the reader to read materials, such as tax forms, that are necessary from a practical point of view. The utilitarian goal is a subvariety of the cultural reproduction goal; in both views the reader is seen as an object, and strategies for teaching reading are likely to involve manipulation of the text.

The third, interactive, view sees the reader as subject. There are various orientations within this viewpoint, such as the romantic (exemplified by bibliotherapy), the cognitive, and the social (as in the work of Kozol, 1978, and Freire, 1970). In the social interactive

framework, the basic question to ask in evaluating the reader's ability to read a text is: Do people at Level Y get the judgment that they have read this text?

Reader-as-Subject vs. Reader-as-Object Approaches

As an example of the difference in approach between social interactive and reader-as-object models, consider a text such as:

(1) Through the years the nation Japan has trained the skilled workers needed in those factories. The result is a prosperous industrial nation.

A typical criterion question that might be asked of a reader in a reader-as-object model is:

(2) Is Japan often thought of as a well-to-do successful nation?

Freire (1970) would argue that a question of this type doesn't provide a very full criterion of understanding. The relevant question, in a social interactive view, is: What does it really mean to read a given text? Do you want people simply to be able to reproduce information, or do you want them to have something like social interaction with the text? Hence, the social interactive model would ask the examiner a question like (3):

(3) Is a question like (2) a measure of whether the passage is readable to the answerer?

Readability formulas may predict the readability of a passage for many kids, but other factors are involved, and it isn't clear

that writing a better readability formula will make any difference. What makes a passage easy to read is not necessarily text-derived. Using a number of factors, it is possible to make a text easier, but how do we know whether the text requires the same task after this has been done? Why, for example, do we assume that reading the "duck" story is the same task as reading something like "Wally's Story?" How do we know that in making reading easier we haven't made it a different task than what we want children to eventually be able to do?

Cultural Attitude and Purpose

The ability of children to read a text interacts with cultural values and with the purposes for reading the text. For example, in one first-grade group of children, the children were able to read a complicated menu containing idiosyncratic spellings and abbreviations in order to find out what was for lunch on Wednesday, but could not read a "simple" text that said only "dodo bird" (under a picture of a dodo bird). In the latter case, the children wanted to read the text as "doo-doo bird," presumably because of the attractiveness of the taboo expression and the orthographic analogy with the word do, even though they knew phonics principles that should have given them the correct pronunciation. In another example, an 8-year-old confronted with the sentence It was three long hours became so angry that he refused to continue reading the text on the grounds that three

hours is three hours. In cases such as these the "readability" of the text depends on factors of cultural attitude and purpose, and not on the length of words and sentences.

Model of Differences

The overall problem we are dealing with in trying to describe readability is: How does a model of text features fit the material of texts? This should not be seen as a relative ranking question, for which it is satisfactory if a model of differences can be devised that fits some comparative ranking. Rather, we need to ask: Can we come up with a model of differences that fits all relative rankings of passages?

Factors Not Considered by the Formulas

On the Flesch scale (1948), Graham Greene's The Heart of the Matter receives a rating of 86, and Faulkner's The Sound and the Fury 97, indicating that these texts are very easy. In general, the readability scales often give strange and counterintuitive ratings.

There do seem to be some ways of patching things up. For example, in looking at a text, the following factors could be carefully considered in addition to the score on a readability formula: (a) degree of abstractness, (b) usage of metaphor, (c) unencoded inferences, and (d) sophistication of topic.

Word Familiarity and Context

There are some specific problems that occur even with the more sophisticated kinds of formulas such as t-unit analysis (Hunt, 1965) and Botel and Granowsky's (1972) syntactic complexity formula. These include the question of what is meant by "word familiarity." The idea that there is a single measure that can be called "word familiarity" ignores the questions of familiarity in oral vs. written language, background of the reader, and the context in which the word is used. Compare, for example, (4a) with (4b):

(4a) Then he looked back, and to his great joy he perceived that the Lillian (a boat) had gone off in a wild direction and was yawing all over the pond.

(4b) When a boat is in a race it must be careful not to yaw.

In (4a) there is sufficient context to enable the reader to figure out the meaning of the word yaw, while in (4b) it is quite difficult. Readability formulas that include a measure of word familiarity are not sensitive to such differences, and would rate yaw in (4a) as just as difficult as yaw in (4b).

Context may also enable the reader to understand a difficult word via a self-defining structure, as in (5):

(5) Beetles are one of the few kinds of insects that make their own light. Many glow only while they are larvae, the young wormlike form.

In (5), the text defines larvae, but a readability formula would not take account of this fact.

Similarly, context may make words that should be "familiar" (and therefore easy) more difficult, as in metaphorical usages. In a text like

> (6) . . . Jan had spoken a declaration of friendship that would make other white men hate him. A particle of white rock had detached itself from that looming mountain of white hate and had rolled down the slope, stopping still at his feet,

the words rolled down the slope, for example, are made more difficult by the fact that they are used metaphorically. Again, measures of word familiarity discount this sort of fact.

Syntactic Complexity

Measures of syntactic complexity are similarly inadequate for capturing the real level of difficulty of text for a reader. One problem is that the number of ways to classify any given sentence is often extensive; there is likely to be little agreement among classifiers. For example, the distinction between headless relatives and embedded questions may be very difficult to draw.

Freeman, Larkin, Humphrey, and Yanofsky (1978) point out that judgments of syntactic complexity must be made in terms of some theory of how language works, and cannot be determined simply on the

basis of surface structural configurations. An example is the differences among Joe's lucky brother, Joe's lucky day, and Joe's lucky penny. The relationships among Joe, lucky, and the noun following lucky are different in the three cases, but surface structure hides this underlying complexity, while a syntactic complexity scale will treat the three as identical structures. Some other examples of this type of problem: We expect a good team to play is ambiguous, but possible differences in complexity between the readings will be ignored by the formulas. A similar situation is the ambiguity in some cases of embedded questions and headless relatives, as in Marie wondered what Sally did, where there may be real differences in reading difficulty between the two readings, "Marie wondered what it was that Sally did" and "Marie wondered the same thing that Sally did."

In some cases measures of syntactic complexity seem to overpredict difficulty, as in (7a):

(7a) The policeman who is usually on the corner wasn't there today.

(7b) The policeman usually on the corner wasn't there today.

Note that (7a) contains a full second clause, whereas (7b) contains a reduced relative clause (via whiz deletion). The Syntactic Complexity Formula (SCF) discussed in Granowsky and Botel (1974) assigns a complexity count of two to "dependent clauses" but only one to adjectives and prepositional phrases. Thus the SCF judges (7b) to be easier than

(7a). Based on the work of Fodor and Garrett (1967) and others, it seems likely that (7a) is actually easier to process, since it provides more clues to the underlying structure, making grammatical relations more explicit. So it appears that such a measure of syntactic complexity as the SCF, as well as one based on sentence-length, will make the wrong prediction here.

A somewhat similar situation occurs in regard to passives, which will automatically be rated by formulas as more difficult than actives. While it may be true that passives are harder in general, actives are clearly more difficult in contexts which topicalize the underlying subject. This accounts for the strangeness of (8b) as a TV announcement:

(8a) Tarzan will not be seen today because of the following presentation.

(8b) You will not see Tarzan today because of the following presentation.

(Freeman et al., 1978)

Clearly, the passive is better because Tarzan is in topic position, whereas (8b) seems to be stressing the viewer. Formulas which automatically consider passives to be harder to understand than actives, regardless of context, will often make wrong predictions.

There is also a problem with "simple sentences," as illustrated by the contrast between (8a) and (8b):

(8a) I like to telephone.
I want to telephone Sam.
I dial his number.
Sam's telephone rings.
I hear a noise in my telephone.
The noise tells me Sam's telephone is ringing.

(8b) I like to telephone. I want to telephone Sam, so I dial his number. Then a noise in my telephone tells me that Sam's telephone is ringing.

As pointed out by Freeman et al. (1978), the use of separate sentences suggests that each fact is important in and of itself and leaves the relationships among the facts unspecified. Measures of complexity, however, would rate (8a) as much simpler than (8b).

Leveling-Out

Overall, readability formulas do not seem to be adequate as measures of difficulty. On a more general level, it may be that the notion of "level" (such as "level" specified by a formula) is not useful in the teaching of reading. Nor does the use of the concept "level" seem to be very useful when applied to grouping children according to ability for the purpose of reading instruction; the result of such grouping is that the gap between lower-group and higher-group children widens. It may be that "leveling-out," which doesn't work in grouping children, really can't be expected to work in evaluating texts either.

This view is supported by the work of Hess and Takanishi (1974), who have found that properties of classroom materials do not have much

impact on the amount of time spent on text, a measure which is correlated with comprehension of the text. Rather, interaction with the teacher seems to be the important factor.

Research suggests that the reading level of a text has very little to do with how well a child reads or how quickly the child learns to read. This is in direct contradiction to the learning-to-read theory which suggests a hierarchy of easiness to hardness on the basis of texts. The questions which must be addressed if we are to develop a theory of the reading task are: How important are texts? Are characteristics used by the readability formulas correct? Are sentence length and word familiarity hierarchical features? Does the task of reading have to do with length and familiarity? If we answer this last question in the affirmative, how do we know that this is an appropriate definition of the task of reading? There is a confusing blend of text-specific and generalizable characteristics in the learning-to-read theory, such as the assumption that the ability to read Here is a duck will generalize to the ability to read Here is a lamb. This is not a very safe assumption.

Zone-of-Proximal-Development Tutorial

An alternative way of studying the process of learning to read is the zone-of-proximal-development tutorial. The goal of this type of study, as exemplified in the work of Wolf (1976), McKenzie (1977), and Etter (Note 4), is to answer the questions: What goes on in interactions

between text and novice vs. accomplished readers? The answer to this question should reveal the components of the reading task, and should demonstrate what is relatively accessible to the novice, which is the real criterion of easiness.

The value of taking this point of view is demonstrated by studies such as Fereirro's (1978) work with prereaders, in which she used Piagetian methodologies to determine what prereaders think is important in reading. One of her findings is that 5-year-olds do not believe that verbs are written on the page. Luria (1977-78), also dealing with the question of what children expect when they go to the literacy task, has stressed the importance of children's idiosyncratic writing and of discovering what the child thinks is writing. This approach takes as its starting point the question: What do children already know?

Discussion

Charrow: Studies of the literacy demands on government clerical workers show that many not-too-well-educated people claim not to be able to read, but that this is because reading has been defined in a certain way, and in fact in their work these people perform a whole range of tasks that one could recall "reading," including reading whole pieces of text.

Griffin: The zone-of-proximal-development approach would involve sitting down with one of these people and figuring out what can be done

to make them able to do more than they're already doing. Then, when this has been done with a number of people, you can begin to abstract out of your findings a scheme of what you can do with people at different "stages." What you end up with is not a set of levels related by quantitative differences; instead you tend to get a set of functionally different systems.

Green: One potential device for showing the relatedness between sentences when connectives have been deleted is paragraphing, but in primary texts there is a tendency to eliminate this sort of structure as well.

Griffin: Paragraphing is probably not very available to the teacher as a device for showing connectedness because of the tradition of oral reading in which the text tends to be broken in strange places by shift of turn from child to child, perhaps on a sentence-by-sentence basis.

Davison: Although it seems clear that there may be difficulties involved in drawing inferences in cases where there is no explicit connection between sentences, it is difficult to think up an experimental paradigm for studying what goes on during reading when this is the case.

Griffin: There are some research programs that purport to study the differences between texts with and without connectives, but a major problem is that you don't really have the same text when the connectives are in that you do when they are taken out. Apparently sometimes having

only paratactics is harder, and sometimes having only syntactics is harder, and sometimes changing from paratactic to syntactic changes the meaning of the text.

Osborn: Given the state of our knowledge, it's very important that all steps toward improving reading instruction be taken with a heavy dose of tryout in classrooms. For example, even though (5) contains a defining structure, children may not be acquainted with the construction and may not realize that larvae is being defined.

Griffin: This is the sort of thing that would be found out in a tutorial.

Osborn: But we cannot count on teachers being sensitive to it in the kinds of classroom settings in which most reading instruction is now being carried out.

Griffin: There are ways of basing an actual curriculum on the zone-of-proximal-development tutorial, such as that described in A. K. Markova's The Teaching and Mastery of Language (1979).

4. Jobs You Shouldn't Count on Readability Formulas To Do

Bertram Bruce and Andee Rubin

It is easy to caricature readability formulas--to find a case where a readability formula rating is empirically invalid and counter-intuitive. Without appropriate analysis, however, such an example does not show why the readability formula fails nor does it distinguish among

different situations in which the formula might be more or less appropriate. In this paper we look at some cases where the formulas seem to go awry and put them into a more general framework based on questions like the following: What are the factors that affect readability? Why does one need a measure of readability? What are the assumptions underlying readability formulas?

In considering the factors that affect readability, let us first take a step back and, without any preconceptions, think about the kinds of questions a reader might ask her or himself in determining whether a particular text is difficult to read. Questions such as the following might be relevant:

1. How do I feel? Am I tired? Hungry? Do my eyes hurt? Am I distracted or preoccupied?
2. How interested am I in this topic or this story?
3. What do I already know about the subject? Do I have enough background knowledge?
4. How similar is the writer's language to mine?
5. How plausible to me are the writer's presuppositions? What do I have to take for granted in order to understand this text? Which of my own beliefs must I temporarily lay aside?
6. Why am I reading this? Do I want a clear model of all the facts presented in the article or is a general understanding sufficient? Is my purpose merely escape?

7. How long do I have to read this? How does this limitation affect my reading goals?

8. What do I want to do with the information I get?

This mere beginning of a list reflects the personal, interactive nature of reading. Notice that each question examines the relationship of the text and the reader; none relates to the text in isolation. In this view of reading, it is bizarre to think of a text as having a degree of readability in and of itself, apart from considerations of the reader and the reader's purpose.

In spite of the plausibility and importance of this image of reading and readability, readability formulas which purport to evaluate texts out of the context of the reader-text interaction have continued to flourish. One obvious reason for their growth is that a number of jobs exist for which a simple measure of text difficulty would be enormously useful. These tasks include: designing (writing, selecting, and adapting) texts appropriate to the level of a child in school, choosing among trade books for children; choosing passages for tests, evaluating difficulties in reading; making writing clear for adults, designing materials for special populations; and writing and evaluating materials to be used in research. We want to consider here a few of the jobs for which readability measures are or might be used, specific examples of each, and problems that come up in using the formulas in these ways. (In some cases, readability formulas are not used explicitly, but similar considerations of vocabulary and sentence length are applied less formally.)

For each problem we identify we would like to ask: Is the problem in this example an idiosyncrasy, or does it reveal some flawed assumptions underlying the readability measure? After looking at several examples, we will be in a better position to address wider issues, such as: Are measures of readability useful for some jobs but not others? Are there any jobs for which such measures are useful and appropriate?

Job 1: TV Captioning for the Deaf

Deaf adults in this culture are doubly isolated from mainstream society--not only is it difficult for them to communicate with most other people around them, but they also lack access to the television programs so many people watch. The process of captioning TV shows must take into account two characteristics of deaf adults' language capabilities for which readability formulas could plausibly provide some assistance. First, deaf adults' reading rates are generally lower than those of hearing adults and significantly lower than the rate of oral presentation, so that a simple transcription of the audio portion of a show would not be readable in the time available. Second, deaf adults are believed to have more limited syntactic and vocabulary abilities than hearing adults. Since readability formulas focus on these aspects of text, they are obvious candidates for evaluating captions. The following illustrates what happens when captioning is done with prime consideration given to word choice and sentence construction.

Example. The following texts are two versions of introductory comments to a television documentary on Arab-Israeli relations. The show starts with a picture of Marilyn Berger speaking on location in the Middle East. Throughout the passage, we see the same scene, so virtually no information is communicated in the video portion of the show. The first version below is the captioned text; the second is the original.

(1) [1] I'm Marilyn Berger. [2] I first came to the Middle East 11 years ago after the 6-day war. [3] Many Israelis thought it would be the last war. [4] This program is not about armies or diplomats. [5] It is about 2 families caught in the Middle East conflict. [6] The program was filmed 3 years ago. [7] But the same feelings remain today as they have for the last 30 years. [8] The feelings remain although there was hope from the historic visit to Israel by Egypt's President Anwar Sadat. [9] This program was the idea of an Israeli TV producer and a Harvard professor. [10] They worked with an Egyptian newsman. [11] There is no attempt to decide who is right or wrong. [12] The purpose is to understand the conflict from the point of view of the people who have the most to win or lose. [13] Recently there have been more deaths and violent statements. [14] In the following program we will hear the softer voices which aren't often heard, but which are no less important. [15] One family was filmed in Egypt. [16] The other family was filmed in Israel. [17] Both families have had a great loss--the death of a child. [18] The film team was surprised and pleased by the sympathy that each family had for the other.

(2) I'm Marilyn Berger. I first came to the Middle East 11 years ago, just after the 6-day war--the war so many Israelis thought was the one to end all wars. What you are about to see is the first in a series about the Middle East. It is not about armies, or governments or diplomats. It's about people and families caught in the conflict. The program was filmed in Egypt, in Syria, in Jordan and Lebanon, and here in Israel some 3 years ago. But the same feelings, the same dilemmas persist today, as they did 3 years ago--indeed as they have for the last 30 years. They persist despite Egyptian President Anwar Sadat's historic visit to Israel last fall, and all the hopes that visit aroused here and around the world. Before we begin this series, a word about how it all came to be, and about some of the unusual people who were involved in it. Back in 1973, Zvi Dor-Ner, an Israeli television producer, and Professor Roger Fisher, of Harvard, who shared a concern for the Middle East and an interest in television had an idea. By July 1974, the idea had taken shape in a series of broadcasts. Fisher and Dor-Ner became part of a team that included Professor Nadav Safran of Harvard, a leading Middle East scholar, Mohammed Salmawy, an Egyptian newspaper man, and a group from WGBH in Boston. Each program in their series is devoted to a specific aspect of the Middle East conflict. It does not try to resolve who is right or wrong, if indeed there is a right and a wrong. The purpose was then, and is now, to try to see the conflict through the eyes of the people who have most at stake; to see it as they see it. Recently there have been new casualities, more deaths--soldiers and civilians. There have been more of the kind of violent and extreme statements that are so much a part of the dialogue in this part of the world. In these programs you will hear something different--those softer

> voices that may be drowned out so easily, but which are no less important to listen to. The first broadcast in the series that you will see tonight centers on two bereaved families--one Egyptian, one Israeli--following the greatest loss a parent can suffer, the loss of a child. When Zvi Dor Ner, filming in Israel, and Mohammed Salmawy, filming in Egypt, compared their separate efforts, both were surprised and pleased. The sympathy each family had for the other was something neither man had expected.[3]

The readability level of the original text is 11.4, that of the adapted text 7.4 on the Fog readability scale (Gunning, 1964).

Problems. The problems with this adaptation arise mainly because the higher-level discourse structure of the text was neglected in its construction. Thus, although individual sentences are "simplified," the overall text is less coherent. For example, the deletion of the repeated phrase "the war" from the original text causes a shift in focus in the adapted text (between sentences [2] and [3]) to the Israelis' beliefs. The result is a garden path reading of sentence [3] which at first suggests "Many Israelis thought it would be Marilyn Berger's last visit." In the original, the focus is maintained.

A similar problem is the rough transition between Anwar Sadat (sentence [8]) and This program (sentence [9]). Organizational material from the original ("Before we begin the series, a word . . .") has been deleted, leaving it to the reader to forge some connection between Anwar Sadat and the program.

Third, in the adapted version it is unclear who is not attempting to decide in There is no attempt to decide who is right or wrong (sentence [11]). Is it the producer, the professor, the newsman, or Anwar Sadat? The original contained sufficient context to make the answer clear.

Finally, the poor connection between the last two sentences ([17] and [18]) tends to lead the reader to the unfortunate initial impression that the film team was surprised and pleased by the death of the children. In this case the original text also has the potential of being misunderstood, but the adaptation has made the problem worse by deleting the mention of "efforts."

In general, the problems with this selection appear to follow partly from the rate constraints imposed by captioners. Captions are written in general for a reading rate of 120 words per minute regardless of the content of the caption or the video. Since in this particular case the screen showed only a talking head, one approach might have been to increase the length of captions (and therefore the presentation rate) so that the text would have been more coherent.

Assumptions. One assumption about the use of readability formulas relevant to this job is that they are being applied to "honestly written" material, i.e., material that has not been specifically 'written down" to meet the demands of the readability measure. One reason that measures of vocabulary and syntactic complexity work at all is the strong correlatior

between real conceptual complexity and complex words and long sentences. In trying to adapt material to fit a readability measure, words and sentences may be changed, but conceptual complexity remains constant. Thus, a formula becomes an inaccurate reflection of the passage difficulty.

Another assumption of the adaptation is that higher-level structure can be ignored. What is most often deleted is organizing material and connectives, resulting in the kinds of problems we saw above.[4]

Job 2: Elementary-School Trade Books

Trade books written for elementary school children are often graded via readability formulas so that they can be matched by teachers, librarians, and parents with the reading abilities of children. Even though it might be argued that young children have very individual tastes, these readability ratings are made on a uniform basis, focusing on vocabulary and sentence length. Some readability formula advocates might even argue that word-level considerations are of primary importance for young children whose reading vocabulary is sharply limited. As just one example of the pitfalls of this approach, let us look at one trade book which, according to readability formulas, should be easy to read.

Example. Don't Forget the Bacon (Hutchins, 1976) is an elementary school trade book which scores between grades 1 and 2 on the Fry (1968) scale. The book concerns a little boy whose mother is sending him on a

shopping errand. On the first page the reader is shown the mother holding a basket and coin purse and saying to the little boy (via a cartoon-type balloon), "Six farm eggs, a cake for tea, a pound of pears, and don't forget the bacon." On the next page the little boy is pictured carrying the basket and coin purse, walking along a street where there are three fat people standing on the corner, and thinking (via a cloud figure), "Six fat legs, a cake for tea, a pound of pears, and don't forget the bacon." The book proceeds in this fashion, with the text consisting only of the little boy's thoughts and a few similarly structured instances of speaking to the storekeeper and to his mother.

Problems. The general impression most adults get upon reading this book is that it is too difficult for early elementary school children. In addition, we have interviewed several third and fourth graders who found the book (but not the individual words) difficult. One reason for the discrepancy is that much of the information in the book is communicated in the pictures. Even the fact, for instance, that the little boy is walking down the street is not explicitly stated, and the various scenes which change his memory of what to buy are pictured, not described.

Second, enjoying the book relies heavily on appreciating its symmetrical structure. The little boy forgets the items on his list one by one by going through a series of transformations of the linguistic material he is trying to remember, then remembers the items in the opposite order, until he arrives back at the original list.

A third aspect of the book not captured by readability measures is the important distinction between thought and speech, indicated by the standard balloons and clouds used in cartoons.

Finally, even though the book would not be considered poetry, an appreciation of rhyme and how people use it as a memory aid is an important component of comprehending this text. Thus, the potential sources of difficulty for the book are not attributable to its word difficulty or sentence complexity, yet these are the only factors that readability formulas consider.

Assumptions. The problems discussed above illustrate the invalidity of the assumption that it is sufficient to evaluate the readability of a text in terms of lexical and syntactic factors alone, ignoring such factors as pictures, graphic conventions, and rhyme.

Another assumption often made in the use of readability formulas is that it is possible to apply formulas which represent statistically derived averages to individual books and to children. In the case of Don't Forget the Bacon, this assumption is clearly unwarranted, as this book seems generally too difficult for its supposed age range. Such examples will always come up in using any statistically-derived quantity as a predictor in an individual case; this fact urges caution in this common use of readability formulas.

Job 3: Reading Comprehension Tests

Reading comprehension tests are another domain in which written material must be graded or scaled. Although we are not sure what part

readability formulas play in the construction of reading tests, we suspect that either the formulas themselves or related considerations of vocabulary and sentence difficulty come into play in the screening, selection, ordering, and norming process. The following example also illustrates some assumptions about the process of reading and comprehending texts which are common to both reading tests and readability formulas.

Example. The following is a passage from the ETS Cooperative English Test of Reading Comprehension, 1960:

> As to clever people's hating each other, I think a little extra talent does sometimes make people jealous. They become irritated by perpetual attempts and failures, and it hurts their tempers and dispositions. Unpretending mediocrity is good, and genius is glorious; but a weak flavor of genius in an essentially common person is detestable. It spoils the grand neutrality of a commonplace character, as the rinsings of an unwashed wineglass spoil a draught of fair water. No wonder the poor fellow who belongs to this class of slightly favored mediocrities is puzzled and vexed by the strange sight of a dozen men of high capacity working and playing together in harmony. (p. 7)

One of the comprehension questions asked about this passage is the following:

> The writer suggests that persons of exceptionally great ability
>
> (a) tend to like and appreciate one another.
>
> (b) dislike the company of ordinary men.

(c) are likely to be jealous of one another.

(d) are essentially common except for their genius.

Problems. Although the writers of the test consider (a) to be the correct answer, some people choose (and can justify) (c) as their answer. People who select (c) interpret the first sentence to mean that clever people are jealous of one another (the same interpretation those who choose (a) make), but then equate "clever people" with "geniuses." Although there are other indications in the passage that the author does not consider clever people to be geniuses, this evidence does not change their opinion, so they assert that "persons of exceptionally great ability are likely to be jealous of one another." In our admittedly limited sample, those who chose (c) considered themselves members of the lowest of the three groups of people identified by the author (ordinary people, slightly gifted people, and geniuses); one is tempted to conclude that they were therefore less conscious of the intended differences between geniuses and clever people. Even if this interpretation is incorrect, it points out how quite personal differences (in this case, in self-image) might affect comprehension.

Assumptions. Users of readability formulas must assume that the intended reader and the rater of a passage do not differ on whatever measures (vocabulary, life experience, purpose for reading) are relevant to the evaluation of the text. In the case of a reading comprehension test, the "rater" may be a person, a committee,or a norming process; with

readability formulas, rater characteristics are derived from the population on which texts are graded, and are reflected, for example, in the choice of words for vocabulary lists. Where differences exist (as they do most obviously here and in a later example), the use of readability formulas (or standardized comprehension tests) is less justified.

This passage illustrates for a second time the importance of honestly written material. Since this material was obviously not written to be excerpted, but was intended as part of a larger text, this method of constructing passages for tests violates the assumption of honestly written material.

Finally, it seems clear that the assumption that language rather than conceptual content is the main determiner of text difficulty would not apply to this passage; the conceptual content is quite complex and could probably not be made any easier (in fact, might be made more difficult) through such procedures as shortening sentences and simplifying syntax.

Job 4: Remedial Reading Texts

Junior-high and high-school age students who have trouble in reading present a difficult problem: Their reading skills may be inadequate for the texts directed at their age level, while texts which they could read easily are typically directed toward much younger children and may be boring or embarrassing for them to read. One solution to this problem

is to identify a set of texts that score low on readability formulas but have high interest value for the older student.

There are several criticisms to be made of the high-interest/low-readability-score solution. First, it is extraordinarily difficult to select high interest texts for an individual student, especially one who is not much of a reader. Giving the student the power to choose (and reject) texts might be more effective and just as pedagogically sound. Second, it might be more useful to search for a meaningful text and task for which the student would be willing to invest some effort. Third, a low readability score is no guarantee of true ease of reading. This is particularly the case for texts constructed or adapted to fit a readability formula.

Nevertheless, readability formulas or their kin appear to be widely used in designing and choosing remedial reading texts. The two examples discussed below illustrate some of the factors this approach fails to cope with. Both passages score about grade 5 on the Fog readability formula.

Example 1. The first example, "Indian Occupation," is from Clue Magazine, No. 2 (Education Progress Corporation, 1972, 1979).

> The Indians had not heard from the government. The suit for Alcatraz was still not settled. The Indians were discouraged and angry. They did not know if their goal could be reached. Some people wanted to tear down the buildings. "The White Man is our foe," they said. "He took our land 300 years ago. It's true! The White Man wrote treaties, but they were all a hoax."

Other Indians said, "Wait! We must build a place here that we can boast about. We must have a school. It's dangerous for our children to roam through these old buildings. We need food, too. We must hoe the soil and plant tomatoes, potatoes, and fruit."

Suddenly someone roared, "Fire! Fire!" A fire had started in an old building. Unfortunately, the boards made good fuel. The flames soared high. There was no water to soak the buildings. The only water on Alcatraz was the drinking water brought by the boats. The Indians had no pumps to bring water out of the bay. Finally, the roaring fire was reduced to coals and burned itself out. There were no clues to tell how it started.

Other problems came up. Food and water did not come when they were due. Boats cruised by, but they didn't stop at the island. Some of the Indians began to loaf. They forgot about their oath to work together. Richard Oakes decided to leave the island. Others said, "If he goes, we'll go too." Nobody could coax them to stay.

Problems. There are numerous problems with the passage, most of which should be obvious on a careful reading. To mention just one: Note the bizarre introduction of Richard Oakes at the end of the story; he turns out to be the central character, but has never been mentioned previously.

Part of the difficulty with this text probably results from the attempt to maintain a particular readability level, and part from the effort to introduce the written forms of particular sounds (long o and u). Maintaining a coherent story line clearly was of secondary importance.

Assumptions. One assumption this use of readability formulas illustrates is that one can afford to ignore higher-level structure in assessing the readability of a text. This assumption is, of course, violated by texts which have no real coherence, such as the preceding one.

A second assumption is that a text need not be honestly written. This material was written just to do the teaching function, and thus does not qualify as "honestly written."

Example 2. A second remedial reading passage intended for older children is "Shine" from the Bridge series (Houghton Mifflin, 1977; see Simpkins, Holt, & Simpkins, 1977). The following is an excerpt:

> Shine was a stoker on the Titanic. The Brother, he shovel coal into the ship furnace to make the engines go. Now dig. Check what went down on the day the Titanic sunk. Shine kept on going up to the captain of the ship. He kept on telling the captain that the ship was leaking.
>
> Shine run on up to the captain and say, "Captain, Captain, I was down in the hole looking for something to eat. And you know what? The water rose above my feet." (pp. 1-4)

Problems. While the text from which this passage is taken does seem to be better as a story than the preceding text, it poses some problems for white students because it builds on oral language, in particular on Black English Vernacular (BEV). Some misinterpretations white students made when reading passages from the Bridge series were (a) to read brother as meaning male sibling, (b) to read Russ say,

". . ." as Someone made/is making Russ say ". . .," and (c) to fail to understand such expressions as they wheels (Standard English, their car). The use of imperatives directed to the reader, such as Now dig, may be characteristic of black folk tales or BEV, but this is unlikely to be familiar to the white student.

Assumptions. The assumption of readability measures that it is possible to gauge the readability of a text without making reference to the cultural background and values of the reader is clearly violated by passages of this type. Readability measures would not tap the white student's lack of familiarity with Black English and the conventions of black folk tales. Conversely, of course, a student who was to some degree alienated or removed from white middle-class culture would have analogous reading difficulties that would not be accounted for by readability formulas. A major determinant of true readability is the match between the cultural attitudes, beliefs, and values of the author and those of the reader.

Job 5: Basal Readers

Basal readers are texts whose main purpose is to be used in teaching reading. Most of the basal readers feature a "controlled vocabulary," which permits only a slow, paced introduction of new words. They also exhibit similar controls over sentence length and syntactic complexity. The standards for language of basal readers define implicit readability

formulas, but many of the series apply explicit formulas as well, in either writing or selection of passages (Rob Tierney, personal communication).

Example. The following is an example from a second-grade basal reader (Ginn 720, Level 6, Unit 3; see Clymer, Wang, & Benedict, 1976):

"See the sights!" called the tall man.
Every day the tall man came to 5th Street.
Every day he called, "See the sights! See the sights!"
One day Dan was walking on 5th Street.
The tall man was there.
He was calling, "See the sights! See the sights!"
Dan saw a big sight-seeing bus stop on 5th Street.
There was a sign on the back of the bus.
The sign said, "See the Sights! See the Sights!"

"What sights?
Where does that bus go?" Dan thought.
The next day Dan walked up to the tall man.
"I want to see the sights," he said.
"When can I take the bus?"
"You're too late today," said the man.
"Come back next Saturday.
Saturday you can see the sights!" (pp. 89-91)

Problems. This selection lacks coherence, even though it is supposedly easy to read according to readability formulas. In the process of controlling vocabulary, sentence length, syntactic complexity, and so on, the basal reader authors have had to ignore other crucial characteristics

of the text. As a result, a child who looks for a familiar structure, for example, conflict and resolution, will find a story which violates that expectation. Laying aside for a moment the question of the aesthetic or pleasure value of such a story, one might simply assume that the story is harder to read because of its inadequacies in terms of characteristics such as conflict, suspense, surprise, and humor. Children fed a steady diet of basal reader stories may develop a conception of stories which discourages them from exploring other texts and which does not match the passages they encounter elsewhere.

Steinberg and Bruce (1980) reported on a study of story characteristics that is relevant to this problem. They coded stories for rhetorical structure, point of view, conflict, amount of insight into characters' thoughts and feelings, etc. One of their findings was a dramatic shift from lower-primary-level basal stories to upper-primary-level stories in the amount of inside view (insight into characters' thoughts and feelings). A similar situation obtains with regard to interpersonal or internal conflict; the upper primary basals were rated much higher on this parameter. Again, there is a major difference between the kinds of educational reading materials children receive in lower primary grades and those they receive in upper primary grades. Perhaps this difference reflects the fact that the lower level stories are constructed to conform to readability formula constraints, whereas the upper-level stories are typically just selected and adapted from trade books.

Assumptions. The basal selections show again that higher-level factors such as conflict and inside view cannot be ignored in assessing readability. It may well be that children who have not learned in basal texts to understand patterns of conflict, etc., in stories may not be adept at dealing with these factors in upper-level texts.

Assumptions About Readability Measures and Their Use

The preceding examples have illustrated various ways in which readability formulas give faulty predictions or even lead to the writing of passages which are harder to read. In each case, one can point to an assumption about the use of the formulas which has been violated. We are led to the conclusion that the formulas are valid only if certain conditions hold. Our list of assumptions has arisen from examination of cases where the formulas have failed, but similar lists have been put forth by designers of the formulas themselves. For example, explanatory material put out by the publishers of the Raygor Readability Estimator states quite accurately some of the limitations of readability formulas:

> Reader interest level, reader experience, or any other personal or ethnic variables are not measured by this or any other estimators of readability. Readability estimators do not measure style or syntax.
>
> Making materials less difficult by shortening sentences and substituting shorter or more common words for longer and more difficult sentences and words may not, in fact,

> reduce the difficulty level indicated when the formula is applied to the new material. The new material may appear easier and show a lower grade level with the estimate, but the concept level may still be high. Readability estimates use variables that predict but do not necessarily control the difficulty of the material. Estimates work best on discursive or narrative prose. Applying estimates to poetry, test items, or other types of nonprose material may produce inaccurate results.

These cautions seem clear enough, and examples such as the ones presented in this paper give strong evidence that the cautions should be observed. Nevertheless, it appears that not only some, but nearly all, uses of readability formulas violate the basic assumptions on their applicability. The problem is that the assumptions restrict readability formula use to trivial cases of little import for educational or social policy. Our examples have indicated that readability formulas should be used only where the following criteria are met:

1. Material may be freely read. Material like captioning for the deaf, which appears on the screen and then disappears after a certain amount of time, cannot be freely read. The time spent on it is limited by external factors, not the choice of the reader.

2. Text is honestly written. The formulas assume that material is not written to satisfy the readability formulas, but rather to satisfy some other communicative goal.

3. Higher-level text structures are irrelevant. The formulas assume that organizational material, information about intentions, goals, etc., need not be specifically taken into account.

4. Purpose in reading is irrelevant. Skimming, test-taking, reading for pleasure, and so on are all taken to be equivalent in determining the readability of a passage.

5. Statistical averages are meaningful in individual cases. Use of the formulas implies that statistical averages regarding both texts and readers can provide useful information regarding the appropriateness of an individual text for an individual person.

6. Readers you are interested in are the same as the readers on whom the readability formula was validated. Any attempt to expand the use of the formula to evaluate materials for readers whose background, dialect, purpose in reading, etc., differs from that of the readers used in validation is likely to lead to difficulties.

Rigorous adherence to these assumptions effectively prevents use of readability formulas for TV captioning, adaptation, selection of texts for readers of different cultural backgrounds, designing special texts for children, selection of text passages, choosing trade books, or designing remedial readers.

We are left with a question: Are there any areas in which the assumptions about the readability formulas are satisfied and the formulas improve on intuitive estimates of the readability of the text? We think

not. The real factors that affect readability are things like the background knowledge of the reader relative to the knowledge presumed by the writer, the purpose of the reader relative to the purpose of the writer, and the purpose of the person who is presenting the text to the reader. These factors cannot be captured in a simple formula and ignoring them may do more harm than good.

Discussion

Tierney: Publishers of basal readers are very much involved with readability formulas--a frustrating situation for anyone who wants to work with a publisher to produce high-quality basals. The experience of Scott Foresman is typical of the dilemma facing publishers of basal readers. In the early 1970s, Scott Foresman tried out an approach called "learner verification of passages," in which, as someone suggested at this conference, they actually did go out and try out passages on children. However, they found this approach to be very expensive, and found that they did not have the technique necessary to do a reasonable interpretation of this kind of research. At the same time, California was developing a policy of checking every single 100-word passage to make sure it met specified readability requirements. The result was that publishers began to produce basals that were very strictly constrained--e.g., a basal at level 2.1 would contain no 100-word passage that scored above level 2.5 on two readability formulas, with no consideration of standard error in the application of the formulas.

This development went into effect in at least four ways: (a) Potential authors received guidelines for avoiding long sentences, sticking to words on particular word lists, etc. (b) Requests were sent out to authors to write within specified constraints, but authors were often unable or unwilling to do this. When this happened, the text would be adapted, and the author might lose out on royalties if he or she was unwilling to accept the adaptation. (c) Already published texts were adapted to meet readability requirements. (d) Editors sent requests to people who were writing teachers' guides to include material that would support the text. This latter action was an attempt to minimize possible negative effects of fitting materials to formulas, and anyone who wants to criticize the use of readability formulas has to address the fact that the texts do appear in a context including such factors as teacher interaction and illustrations.

Kantor: Would you say this is true for grades 5, 6, and 7, or only for the lower grades?

Tierney: All the grades.

Seidenberg: What is the teacher's part in this? What can the teacher do with what is superficially a poor text?

Tierney: Typically the teacher will spend some time talking about a topic before having the students read a text on it, and perhaps will talk about a few specific purposes for reading the text.

Rubin: I think that's a good point, but what does it do to kids' ideas of what the reading task is about? Reading becomes something you can do only if you spend half an hour preparing for it.

Hermon (Linguistics, University of Illinois): Why rely on the individual teacher?

Tierney: I don't mean to say that I approve of the approach I've just described; I'm merely telling you the kind of justification that basal publishers tend to give.

Selden: In the publishing houses I've dealt with, I've encountered less stringency about readability levels. How strong are the prescriptions that go out to story authors? Is there some possibility of honest texts?

Tierney: As you move up in grades, freedom greatly increases. At the lower levels, texts are very tightly constrained, and the editors probably do most of the writing. The real problem with the basals is that publishers are responding to popular sentiment rather than to enlightenment; the only time I've ever seen publishers pay attention to research is when a specific publisher's name is mentioned in the research.

Rubin: The amount of enthusiasm an adult feels for a text is readily communicated to a child, and it seems likely that a teacher who has to wade through three pages of supplementary notes before reading a four-sentence text isn't going to be too excited about the text, and this will be apparent to the child.

Johnston (CSR): When I was teaching, I found some use for readability formulas, namely when a kid would come to me and say, "I don't read very well, but I'd like a book that talks about this kind of thing. What can I read?" It's very easy to criticize readability formulas, but

people who want to get rid of them should propose some other means for doing some of the jobs they do, like helping teachers to locate texts of appropriate reading levels for children.

Griffin: I don't know of any research that shows that matching the level of a text to a child is a better way of teaching kids to read than not. Carol Chomsky's work, in fact, suggests that it is better for a child to read a text that is more syntactically complex than his or her own oral language. In general, it seems that hitting above the level of the learner gives the learner a boost up. Maybe if you had a good text that was very, very hard, the kid could learn to read just as well as if you had a good text that was very, very easy--I don't know, since the research on this has not been done.

Brewer (CSR): That kind of consideration really moves the issue, though. The assumption of this conference has not been that text features don't matter, but rather that readability formulas are not a good way to measure text features, so the question of what to use instead of formulas is a legitimate one.

Johnston: In teachers' college we were taught how to use certain formulas to evaluate the appropriateness of materials.

Zwicky (Columbus, Ohio): Did that work better than reading the text and deciding for yourself how hard it is?

Johnston: I've thought about that, and it's true that people's intuitions are good at that, but people's intuitions have to be built up.

You people who have spent a lot of time studying texts have a good idea of the kinds of things that have an effect, but a lot of the teachers are young teachers who are going out and having to deal with that kind of thing, and need to know what sorts of variables to look out for. Maybe you should look to something like a readability formula as a gross measure, and then sensitize the teachers to kinds of things they might also look at in conjunction with that.

Charrow: I have a better solution than that. Since readability formulas don't work very well, you're really better off giving a child a text you think he might enjoy that may or may not be at his level and then sitting down with the kid and saying, "All right. Did you understand it? Let's talk about it," and in that way having your own ability to judge what's a good text for a given child.

Johnston: That would be a good way to do it, but for a normal range of well-written texts at lower levels, I don't believe that a readability formula is that bad.

Zwicky: I guess we're saying that there are not that many well-written texts.

Johnston: I disagree with that.

Brewer: One line of attack against the pragmatic argument here is what Rob Tierney said--that publishers out there are taking correlation as causation and assuming that the formulas predict comprehensibility, so they go out and destroy the text in order to reach a certain readability level. That's clearly a logical flaw, and a place where we should be involved.

Selden: The publishers don't necessarily think that the formulas predict the difficulty of the adapted text; they think they have to have the numbers in order to sell. They'd probably prefer to sell texts without using the numbers because it would be cheaper to do.

Brewer: But presumably if conferences like this go on and make the argument against them, that will eventually filter into the system.

Tierney: In some states, like Texas, there are state adoption committees that set up rather arbitrary readability requirements, and if textbooks don't meet those requirements, they're off the list. There are some two million children involved, so it's a matter of a large amount of business.

Bruce: We tried to write a program to apply several different formulas to texts, but we ran into terrible problems trying to define what a sentence is. Different formulas seem to have different criteria, and there are lots of unclear cases. It sounds like a trivial question, but it turns out that a large number of examples are hard to classify, and our programmer found that changing the definition of "sentence" changed the readability level of texts by a grade or more.

Tierney: This sort of thing allows the publisher to manipulate the readability of a text by doing things like changing "No, . . ." to "No. . . ."

5. Lowering the Difficulty of Texts Intended for Adults: Implications for "Plain Language" in Legal Documents

Veda Charrow

The study of jury instructions which I discuss here was not originally intended as a test of readability formulas, but the results of the study do suggest that such formulas are not adequate for measuring the extent to which documents such as jury instructions are understandable. The purposes of the study were (a) to look at the "legalese" in codified jury instructions and demonstrate that it causes the instructions to be poorly understood by the average person; (b) to point out the aspects of legalese (grammatical constructions and discourse structure, not just vocabulary or legal concepts) that cause the difficulty; and (c) to demonstrate that if the linguistic problems in jury instructions were removed, comprehension would increase dramatically.

Jury Instructions

Jury instructions are standardized instructions of law set out in form books, and at the end of a trial, the judge reads the appropriate instructions to the jury. If the "wrong words" are used, a case may be appealed; hence, the judge will use exactly the words in the book, even if the instruction is difficult to understand.

The Document Design Project at AIR

The results of the study analyzing jury instructions indicate that grammatical complexity is at the root of difficulties with legalese. However, many psychologists concerned with language (e.g., Kintsch, 1977; Meyer, 1975) have been heading in a different direction. They have stopped looking at grammatical complexity, and are instead concentrating on such notions as propositional density (the number of propositions--in the form of predications--in a sentence). However, the case of jury instructions shows that the study of grammatical complexity is still very necessary, and that there is a great deal more to be learned about what is a grammatically complex sentence.

The Document Design Project at American Institute for Research (AIR) was set up in 1978 under a contract from NIE to find out what makes legal and bureaucratic texts difficult to read, and to help government agencies rewrite official documents so that they are more readable. My colleagues at AIR and I have found that the problem of legalese and bureaucratese cuts across department lines; there is a general problem involving grammatical complexity in the texts that are produced by government offices. And we have further found that many people who can read quite well cannot read bureaucratic documents because they have trouble with bureaucratic language, which may include various sorts of jargon and complex sentence constructions.

The Preliminary Phase of the Study

The jury instruction study, which took place from 1976 through 1978, involved looking at jury instructions as discourse, outside a trial situation. We first chose 52 California civil jury instructions and sent them to trial attorneys, who were asked to rate them for difficulty on a scale of 1 to 11. Half of the attorneys were directed to disregard the complexity of the language, and the other half were told to take into account both language complexity and legal complexity. It proved to be impossible to separate language from legal complexity; the results were very similar for the two groups, and there was general agreement among the lawyers as to what was complex. We then selected fourteen instructions that would constitute a set (opening, middle, and closing instructions), made up a set of facts regarding a case, and tape recorded the jury instructions so that each one occurred twice in succession.

The Task

We received permission to use as subjects people who had been called for jury duty in Prince Georges County, Maryland. The subjects, in a one-to-one interview situation, were asked to explain the instructions by repeating them in their own words. A subject was first given a picture of a fictitious accident and a written description of the "facts" of the accident (to serve as a context for the task). The subject then heard each jury instruction twice, and after the second playing of

an instruction the subject paraphrased it on another tape. All subjects were able to perform the task, but the degree of success correlated highly with a subject's educational level.

Scoring the Results

Instructions were then broken down into smaller idea units (clauses, phrases) for scoring. Subjects' paraphrases were transcribed and compared to the breakdowns of the instructions. An example of an instruction divided into units is as follows:

1. It is my duty
2. to instruct you in the law
3. that applies to this case
4. and you must follow the law
5. as I state it to you.
6. As jurors
7. it is your exclusive duty
8. to decide all questions of fact
9. submitted to you.

To determine comprehension of these units, we noted whether a subject had paraphrased it correctly, omitted it, or stated it incorrectly.

Interpreting the Results

Inability to paraphrase often indicated a failure to understand, but it could reflect other factors as well, such as memory overload, or the triviality of a given idea. We had reliable strategies for discerning the probable reason for omissions in the paraphrases. The following were some apparent sources of difficulty in understanding the instructions:

1. Unfamiliar vocabulary, as in "A proximate cause of an injury is a cause which, in natural and continuous sequence, produces the injury, and without which the injury would not have occurred." Some subjects substituted estimated or approximate for proximate.

2. Unusual placement of phrases, as in the previous example. Some subjects explained the proximate cause as one cause in a sequence, apparently misunderstanding because of the unusual placement of the phrase "in natural and continuous sequence." (In reality, the proximate cause starts the sequence.)

Another example: Given "If in these instructions any rule, direction or idea is repeated or stated . . .," people tended to paraphrase it as "If these instructions are repeated . . .," probably because the phrase in these instructions occurs in a position where a grammatical subject would have been expected.

3. As to constructions. One instruction used six as to's in close proximity; these created vagueness in meaning, and should be replaced with with regard to, about, concerning, etc.

4. Multiple negatives, as in "Innocent misrecollection is not uncommon. . . ." Subjects often got lost amid the negatives and misinterpreted a positive meaning for a negative, and vice-versa.

5. Nominalizations, as in "after a consideration of the case with the other jurors" instead of "after considering/you've considered the case with the other jurors." Subjects' errors increased around such

constructions, as they had trouble discerning who did what. Nominalizations appear to characterize bureaucratic documents, as well. In fact, in some bureaucratic documents there are hardly any verbs at all, apart from is, necessitates, and facilitates.

6. Strings of attributes, as in "a witness who has special knowledge, skill, experience, training, or education in a particular science, profession, or occupation. . . ." Subjects remembered only one or two items in a string, usually the most general ones.

7. Some passives, namely passives in relative clauses where WHIZ-deletion has occurred; e.g., "any insinuation suggested by a question asked a witness," where 'which is" has been deleted (whence the term "WHIZ-deletion") between insinuation and suggested and between question and asked; for this type of structure, 80% of the paraphrases were wrong (as compared with an overall average of 50%).

8. Discourse problems, such as repeating the same instruction in different words when there was no obvious reason for doing so, e.g.:

> "A plaintiff who is not contributorily negligent, and who did not assume the risk of harm, and who is injured as a proximate result of some negligent conduct on the part of a defendant, is entitled to recover compensation for such injury from that defendant.
>
> Thus, the plaintiff is entitled to a verdict in this case if you find, in accordance with my instructions, 1) that defendant was negligent, 2) that such negligence was a proximate cause of injury to the plaintiff, 3) that the plaintiff was not negligent

or, if negligent, did not contribute as a proximate cause to his injury, and 4) that the plaintiff did not assume the risk of harm."

Subjects often thought the second paragraph must say something different from the first; if not, why repeat it?

The Rewritten Instructions

After we had obtained these results, we rewrote the instructions in an attempt to eliminate the difficult constructions and gave them to a second group of 70 people.

Results

Because the general level of education of the second group was much lower than that of the first group, each subject received seven original instructions and seven new ones, so that we could compare performance on original and revised instructions within the second group. We found that paraphrase performance on the rewritten instructions was an average of about 50% higher than on the original instructions. On four of the instructions there was no improvement for the rewritten versions (these had fairly high comprehension scores to begin with), on a few there was a small improvement, on several there was fairly large improvement, and on others there was 100% improvement. These results indicate that one cannot simply assume that his or her rewrite is better than the original; it is necessary to go out and test it.

Applying the Flesch Formula to the Two Sets

We applied the Flesch (1948) readability formula to the two sets of instructions (original and rewritten) and found that for half the instructions, the formula predicted the differences in the right direction, and for the other half it did not. For two instructions that showed large improvement in comprehension, the readability scores were higher, and for one there was no difference. For the four instructions on which there was no difference in performance between the original and rewritten versions, the formula showed better readability. Hence, the readability formula was misleading, as it made the wrong prediction in half the cases.

Conclusion

Contrary to what is usually done in rewriting, some of the changes we made in the instructions involved increasing paragraph and sentence length for such purposes as adding more context. Subordination patterns were also changed (e.g., eliminating left branching and center embedding). Thus, the rewritten instructions may have contained many long sentences (which a formula would rate less "readable"), but the sentence structure was simpler and clearer than the long sentences in the original instructions. (The Document Design Project, 1979, has incorporated several of our findings into its 17 guidelines for language simplification).[5]

Discussion

Seidenberg: Perhaps for lawyers legalese is perfectly clear and comprehensible, so maybe it serves a genuine function.

Charrow: While it would be nice if this were true, apparently it isn't. David Mellinkoff, a law professor at UCLA, claimed in a recent paper (Note 5) that lawyers do not understand legalese very well either. As an example of the difficulties lawyers run into, the word shall is generally favored in legal writing instead of must, leading to ambiguities between the future tense and the notion of obligation or necessity. Legalese is perpetuated because of reusing parts of old contracts . . .

Selden: The legal writing courses that law students take . . .

Brewer: And the desire to maintain a specialized language that non-lawyers will have difficulty understanding.

Kantor: It would seem to be desirable to have positive guidelines rather than simply telling writers what they should not do, since telling them what not to do doesn't necessarily help them do something else instead.

Charrow: Many of the Document Design guidelines are formulated positively--e.g., "Use the active voice" (this particular one is motivated by the need not to leave agents unspecified, rather than by grammatical complexity).

Selden: The Army has found that guidelines per se are not too useful, since writers cannot keep them all in mind while they are writing; it is better to give examples of good writing, along with explanations of particular points.

Charrow: The Document Design Project helps teach good writing by offering courses for bureaucrats.

Some Possible Morals of This Research

Davison: If the researchers had simply used readability formulas, illegitimately translated into guidelines, they would not have been able to achieve the results they did on the jury instruction experiment. Successful rewriting is very difficult and expensive and requires people who know what they are doing--not the mere application of a formula.

Anderson: If the researchers had used only one readability formula to decide whether their rewrites were effective or not, they would have been misled in this case.

Charrow: When dealing with legal language in particular, readability formulas cannot be used because (a) it often is impossible to get away from the type of (legal) vocabulary needed, and (b) long sentences may be necessary, like many of those in the rewrites. If a more adequate formula were devised, probably nobody but a linguist would be able to use it.

6. Readability Formulas and the Adaptation of Texts

Alice Davison and Robert N. Kantor

Readability formulas were never intended by their originators to serve as directions for writing a text to a particular grade level or for simplifying a text so that its level of reading difficulty is lowered.

Writers of readability formulas usually add a disclaimer to this effect (e.g., Klare, 1974-75); and properly so. A distinction should be made between measuring factors correlated in some way with reading difficulty, and actually defining the factors which cause a text to be difficult to read.

But readability formulas involve only a few such factors correlated with difficulty. Once writers know how readability formulas work, it is hard to forget that they involve sentence length and unfamiliarity or length of words, both pervasive features of texts. And even if writers are not told to use short sentences and simple vocabulary, they might independently conclude that doing so would bring the text to the level of reading difficulty which is desired or necessary for the particular purpose for which the text is to be used.

Adaptations and Sentence Length

In this section we want to present some particular cases where evidence internal to the texts in question leads us to believe that readability formulas have had some influence on the writing of materials intended for reading practice. The examples are excerpted from a longer study (Davison, Kantor, Hannah, Hermon, Lutz, & Salizillo, 1980) where we systematically compared the original and the adapted versions of four texts in Parker (1963). There we argue that many of the most appropriate and successful changes could not have been made if the adaptors were using readability formulas as their sole guide in rewriting.

We also argued, as we do here, that there were other changes which were not particularly successful per se, or which in fact made the text harder to understand. These changes did not have any obvious motivations, but since they involved shortening sentences and changing particular words, it is probable that they were made with readability formulas in mind. This feeling is supported by a comparison of the reading levels of the adapted texts with the original versions, as measured by a number of the standard formulas (see Table 1). Put

---

Insert Table 1 about here.

---

crudely, it looks as though certain changes were made to those elements of the text which are measured by readability formulas.

So if the reading level of the texts had to be a certain figure, the adaptor could count on this figure being reached by making sure that there were few very long sentences, that most sentences did not exceed a given length, and that the vocabulary was neither unfamiliar nor complex in length.

Adaptations and Sentence Restructuring

An overall sentence-by-sentence comparison of the original and adapted forms of the four texts studied did in fact show that there were decreases in the average length of sentences as measured in words (rather than syllables--see Table 2). In the adaptation of the longer

texts, a lot of material was simply deleted so that the absolute number of words (and sentences) was decreased (see Table 3). Sentence length

---

Insert Tables 2 and 3 about here.

---

was decreased either by outright deletion of material which we assume the adaptor felt could be left out without distorting the text, or by breaking up longer sentences into their constituent clauses which then become independent sentences. Table 4 shows that the number of long sentences with three or more clauses is substantially reduced in two of the texts studied (two whose original forms were of manageable length for counting clauses by hand). The number of sentences with one or two clauses was somewhat increased. Tables 1 and 2 show that the number of words in the adapted texts were distributed among proportionately more sentences than in the original, so that the sentences were simply made shorter.

Insert Table 4 about here.

One effect of shortening sentences is to introduce uncertainty into the text. Splitting a complex sentence into its component parts often requires that the adaptor delete connective words like subordinate conjunctions or else remove grammatical markers like the infinitive particle to, which are not found in independent sentences. But conjunctions and other clause connectives also have meaning, and deleting them robs the resultant sequence of sentences of some of the author's

original intent, particularly information about the logical relations between the sentences.

Some examples are given below. In (1), the original sentence has an infinitive purpose clause, which is split off from the main clause in the adaptation.

(1) (Disaster in Dayton)

Original: "I'm going down to the contract," said Jack, "to see that everything is all right."

Adapted: "I'm going down to the building project," said Jack. "I have to see if everything is all right."

The altered version is a statement of obligation, not of purpose, which could not be expressed as a main clause without redundancy (e.g., I'm going to see . . .). The expression of obligation might allow the reader to infer what Jack's purpose was in going to the building project, but the change represents a distortion of the text in that the semantic/pragmatic notion of obligation is not described in the original and is not justified in the adaptation.

In example (2), the reader must also infer the correct relation between two clauses in the adapted version:

(2) (California's Giants)

Original: If given a chance before another fire comes, the tree will heal its own wounds by growing new bark over the burned part.

Adapted: If given a chance before another fire comes, the tree will heal its own wounds. It will grow new bark over the burned part.

The reader could infer that growing new bark is the means by which trees heal their own wounds if it is not explicitly stated in the adapted version. Yet an inexperienced reader, or one who does not know very much about trees, might make an incorrect guess and see healing wounds and growing new bark as separate processes, simply ordered in time.

These are quite clear cases of loss of relational information which is explicit in the original and only partially recoverable from the adapted version. Some deletions of conjunctions such as and may simply make less clear which sentences have related topics. But other deletions, while not removing much explicit relational information, may take away hints as to what causes what.

For example, in splitting up the long sentence in the original of (3) into three separate sentences, the adaptor has left fewer clues for relating the clauses. The connection of the first two would convey to the adult reader that the narrator was trying to keep from being afraid through the night, when the dangers were not visible, and nothing much could be done about them until the dawn put a limit to the narrator's conscious effort. The adapted version simply describes a sequence, where the coming of the dawn and looking out are related only in time to keeping up courage.

(3) (Disaster in Dayton)

Original: I had kept my nerve pretty well til dawn, just as the faint light was coming, when we looked out and saw the water whirling by against the bay window.

Adapted: But we all kept our courage up. As the faint light of dawn was coming, we looked out. The water was whirling by.

Splitting up the third and fourth clauses into two separate sentences (which is the result of deleting and saw) makes the connection between the ideas or propositions inferable but not explicit. Clearly, the emotional overtone of the "whole," as expressed by the particular clausal combination in 3 (original) has been obscured.

Deletions of Point of View and Modality

The particular loss of information in this case is also found in many other examples where verbs of perception, speaking, or mental process are deleted. In the adapted version above, the statement The water was whirling by could be correctly attributed to the narrator, since there is no distinction made between the thoughts and perceptions of the narrator and those of the characters in the event. But in the more descriptive texts, there is often a great difference between the ideas of the writer and those of other people mentioned. Learning to distinguish between assertions of the writer and opinions attributed to others, not necessarily shared by the writer, is a major task for an inexperienced reader and should, we believe, be a major instructional goal. However, the adaptations we studied consistently seemed to remove material from sentences which gave information about the source of a statement or which made a statement seem as though it was not asserted categorically by the author.

Information about the source of an idea is deleted when the proposition in question is not disputed by the author, as example (4) illustrates.

(4) (California's Giants)

Original: A railroad freight agent has figured that it would require at least 40 modern flat cars to haul away just the trunk alone.

Adapted: And at least forty freight cars would be needed to haul away just its trunk.

The assertion would be more convincing, it seems to us, if its source is described as a person who would know (perhaps better than the author) exactly what an enormous task such as this would require. Similarly, the example in (4) illustrates the deletion of a description of a source.

(5) (Milk)

Original: Romans were said by Pliny to rub bread soaked in asses' milk on their faces to make them fairer and prevent the growth of beards.

Adapted: The Romans rubbed bread soaked in asses' milk on their faces. They thought that this would make their skin paler. They also thought it would keep their beards from growing!

The assertion about the Romans ought to gain credibility by being attributed to an eyewitness, but this information is deleted, perhaps because the writer would have to describe who Pliny was. (A paraphrase might have been used, such as "a writer who lived in the Roman period . . ."). In the adaptation, the writer adds They thought . . ., to a

proposition which is supposed to be untrue, though in ordinary adult writing, the attribution of a proposition to a source does not indicate that the author disagrees. Instead, the reader must learn to judge in a particular context whether the author agrees with cited soruces, disagrees, or remains neutral.

The deletion of the source of a proposition has several adverse effects, in our opinion. Most important, it obscures the difference between fact and opinion; statements are made boldly, as though they are simply facts emanating from an infallible and authoritative source, the author. Secondly, readers are given no practice in distinguishing shades of opinion, deciding who thinks what, and in evaluating the probable degree of truth which a statement may have. This same effect is found in the deletion of modal information, as in (6):

(6) (Milk)

Original: Nero's wife, Poppaea, took a daily bath in it (=milk) and supposedly had 500 beasts on tap for the purpose.

Adapted: The wife of one emperor of Rome took a milk bath every day . . . She kept five hundred animals to make sure of having enough milk each day.

The adapted version presents the proposition as a known fact, while the original presents it as probable or conjectural, given that it is a fact about a remote period known through possibly unreliable sources.

A third and perhaps less pernicious effect of deletions of source information is simply to deprive the texts of some color or immediacy.

(7) (Disaster in Dayton)

Original: I noticed the water kept rising.

Adapted: The water kept rising.

Since the narrator and the protagonist are the same, there is no chance for the reader to misunderstand the source, yet still the adapted version conveys less vividly the relation between the events and the narrator; in (7), the bare statement that the waters were rising would make the reader less aware than would the original version that the event was likely to make the narrator still more apprehensive of danger.

The net effect of all these deletions is to make the sentences in the adapted version of the text much shorter than the corresponding sentences of the original. The primary motivation to delete this material seems to be to shorten the text, and since the deletions do not radically distort the meaning of the text, the losses of information are considered tolerable by the adaptors. But, as we have demonstrated in the preceding examples, the loss of information in sentence connections may make some of the meaning of the text harder to recover than in the original (examples 1-3). Deletion of information about the source or probable truth value of a proposition, as in (4) - (7), gives the misleading impression that all assertions are equally solid facts and takes away the opportunities for the reader to learn to make judgments about the reliability of an assertion.

Lexical Changes

The changes made in the vocabulary of a text may also lead to loss of information if a more familiar or shorter word is not available as a close counterpart of the unfamiliar word. In the following example, surplus is deleted, in addition to information about quantity, 250 gallons:

(8) (Milk)

Original: In Toronto, a suburban ice-skating rink was flooded with 250 surplus gallons of it (= skim milk).

Adapted: An ice skating rink was flooded with it.

The deletion of surplus takes away information from which the reader might guess why milk was being used in this apparently wasteful manner. A paraphrase to that effect might have been more appropriate: . . . was flooded with milk which was not needed; might have spoiled, etc. But such a paraphrase would have lengthened the sentence and so might have been undesirable if the writer was also under the injunction (tacit or otherwise) to reduce sentence length.

Very often the substitution of shorter or more familiar words must lead to lengthening of sentences. An extreme case is illustrated in (9).

(9) (Milk)

Original: Hippocrates recommended milk as a curative beverage. (7 words)

Adapted: One of the most famous Greek doctors told his patients to drink milk to cure illness. (16 words)

The adapted version is lengthened by nine words to explain who Hippocrates was (saving one word by not mentioning him by name). Curative, which is the least familiar word in the sentence (according to Dale & O'Rourke, 1976) is paraphrased in three words, to cure illness, while recommend and beverage, neither of which is very unfamiliar, are combined as told his patients to drink. The adapted version does convey the information of the original, but in simpler terms and at the cost of lengthening the sentence to more than double its original length.

Many words are given simpler substitutes even though the original words are likely to be familiar to readers of the intended grade level, though perhaps not guaranteed to be familiar to readers reading below grade level. Since the texts we analyzed were intended for students in the seventh through tenth grades who read at the fifth to sixth grade level, it is understandable that all less familiar words might be given substitutes, paraphrased, or deleted in order to remove any possibility of baffling the reader, even if the text itself did not benefit by the change.

It is also interesting to note that the adaptors in some instances went farther than considerations of readability formulas would demand. They carefully changed all obsolete, idiomatic, or colloquial (slang) expressions which might not be in common use, substituting, for example, the term building project for contract in (1) and courage for nerve in (3). They also deleted unfamiliar names, like Pliny in (5), or gave an explanation of who the person was, as for Hippocrates (9) and

Poppaea in the context for (6); she is the wife of a Roman emperor. Such changes show a sensitivity to the difficulties of the text which do not follow from readability formulas, since proper names are not supposed to enter into the calculation of vocabulary complexity in many readability formulas.

Consequences of Formulas as Writing Guides

Using readability formulas as implicit guides to writing or rewriting thus poses the dilemma that in order to simplify vocabulary, the writer is often forced to lengthen sentences; but to decrease difficulty of sentences, the writer must also shorten them. One way to deal with this problem is to delete information, but some of this information may be useful or necessary for interpreting the relations between sentences, and without it, the sense of the text is not preserved. Even if some connective information is left in, the separation into independent sentences creates the effect of a sequence of unrelated assertions. This effect may be seen in the comparison of a sequence of two sentences from the first edition of a text (Bendick & Gallant, 1980) with the same passage from the second edition, where a subordinate clause has been made into an independent sentence:

> 1st edition. You probably saw lily pads, grass, reeds, and water weeds growing in shallow water near the shore. And maybe there were water striders gliding over the surface of the lake, and small fishes darting among the shadows of the lily pads.

2nd edition. You probably saw lily pads, grass, reeds and water weeds. These plants grow in shallow water near the shore. There may have been water striders gliding over the surface of the lake.

(Bendick & Gallant, 1980)

The subordinate modifier growing . . . has been made a separate sentence connected to the preceding sentence by the anaphoric expression these plants. But as a separate sentence, it appears to be an assertion on an equal level in discourse with the preceding, rather than just interesting background information. The next sentence about water striders therefore appears totally irrelevant and inconsequential, because its connection with the description of the lake is less clear than in the original.

Changes such as the ones illustrated in the above examples generally seem to be dictated by the implicit injunctions following from readability formulas rather than by genuine considerations of difficulties within a particular text. We have noted that unmotivated changes lead to loss of information or create possibilities for misunderstanding the original content of the text. It would seem more productive, as we have noted in Davison et al. (1980), to alter the text where its actual properties make it necessary. For example, the relationship between the negative and the rest of the sentence in (11) is fairly difficult to see, because the negative does not just simply negate the verb.

(11) (California's Giants)

Original: When a Big Tree falls, its needle-like leaves do not wither for years.

Adapted: When a big tree falls, it takes years for its needlelike leaves to wither.

That is, the leaves do wither, but not for years.

No readability formula would define this combination as difficult to comprehend. Only a writer using knowledge of language would have defined it as misleading and made the appropriate change.

Reasoned Adaptation

Clearly, an adaptation of a text which is done for the purpose of making it easier to read will involve simplification and substitution, and probably also some reorganization. If this work is done with some sensitivity to properties of texts and to the expressive characteristics of words and sentences, and if the intended audience actually does understand it, it would seem that readability formulas are basically unnecessary. When readability formulas do have influence over writers, we find such influence tends to undo some of the work of adaptors as intelligent writers by causing them to make the text less connected and its meaning less explicit. It seems to us to be a waste of writer's efforts to ask them to simplify a text both according to their own knowledge of language and according to the contradictory injunctions of readability formulas.

7. Comprehension of Captioned Television

Mark Seidenberg[6]

Virtually all of the 13.4 million hearing-impaired Americans (Schein & Delk, 1974) suffer some degree of informational deprivation. For a small percentage, their handicap merely reduces the flexibility and value of their TV viewing rather than excluding them totally from the medium. For most, the consequences are more severe. Little TV programming other than sports can be understood more than superficially without the auditory component. Lipreading provides little additional information because of the small picture size, poor image definition, and unfavorable camera angles. The cultural and informational isolation that results is formidable.

Comprehension of televised information can be increased by the addition of either captions or simultaneous translation of the audio into sign language. Captioning is the favored method both because many hearing-impaired individuals do not know sign language and because, among those who do, captioning is greatly preferred (Norwood, 1976).

The recent development of "closed" captioning via Line 21 of the broadcast signal and its implementation by three networks (Brown, 1980) ensures that the amount of captioned programming will increase greatly. In addition, full-page text transmission such as the Teletext system will soon become available. Thus, an array of innovative electronic text services is likely to be available soon.

It thus becomes important to understand the cognitive and linguistic factors involved in caption comprehension. This paper reports the results of a study of current captioning practices. Transcripts of programs captioned by WGBH-TV (public television in Boston) and videotapes of captioned programs were analyzed along several dimensions. Our preliminary findings suggest that captioning for adult viewers is currently done primarily on an intuitive basis. Little is known about the properties of captioned texts that contribute to their comprehensibility or about the consequences of current practices. In the simplification process, cues to text structure, connectives, and transitional material are frequently deleted. This may result in a text that is syntactically simple but difficult to comprehend because the reader must generate deleted information and integrate different portions of text in the absence of explicit cues to its structure. We have isolated five groups of factors that appear to heavily influence caption comprehension:

1. Display Conventions including the division of captions into lines and successive displays;
2. Cues to Sentence Structure such as relative pronouns and complementizers, which may be retained or deleted;
3. Global Structure Cues such as introductory sentences, transitional words or phrases, and summary material;
4. Lexical Choice including the retention, deletion, or replacement of nonliteral expressions such as metaphors and idioms;

5. Genres of program content; for instance, short and long news pieces, "teaser" stories, features, drama.

These should be the focus of future experimental research in this area.

Previous Research

Existing research indicates that the television viewing habits of hearing and hearing-impaired persons are comparable. For example, an evaluation of the Captioned ABC News performed by the Deafness Research and Training Center at New York University found that among more than 1,100 hearing-impaired respondents, viewing averaged about 3 hours 30 minutes per day. This compares to 3 hours 44 minutes among men in the general population and 4 hours 51 minutes among women. Sendelbaugh (1978) found that hearing teenagers watch about 20 hours of television a week, while hearing-impaired adolescents watch about 30 hours and deaf adolescents about 36 hours.

Little research has been conducted on captioning; most studies have focused on their general effectiveness. Davila (1972), Fischer (1971), Gates (1970), Nix (1971), and Norwood (1976) showed that televised transmission of information to hearing-impaired subjects is improved by the addition of captions. Propp (1972) found that captions were the most effective of four presentation methods (including simultaneous sign language interpretation). A formal evaluation of captioned "Zoom" (Winslow, 1977) found that children viewing captioned programs had consistently higher attention levels than children viewing uncaptioned programs.

There has been some research on display factors. O'Bryan (Note 6) found that a regularly paced flow of medium-length two-line captions seemed most efficient and least fatiguing for deaf readers. Consistent caption placement was also important, and it was found that captions should be shortened over active video scenes and never cover the speaker's mouth. O'Bryan (1975) found that on color broadcasts yellow captions were not as effective as white captions.

In terms of linguistic structure and content, the research is sparse. An evaluation of captioned television for deaf adults (Bickman, Roth, Szoc, Normoyle, Shutterly, & Wallace, 1979) showed inconclusive results regarding the effectiveness of edited versus verbatim captioned scripts. The study suggested that "more research is necessary to examine and fully delineate the important linguistic link between the captions and information transmission."

In sum, previous research has shown (a) that hearing-impaired persons watch television, and (b) that captions are beneficial. However, almost no research exists concerning caption content.

Issues

The goal of the present research is to identify variables that affect caption comprehension in order to develop guidelines for caption writing. Caption comprehension is a language experience that draws upon skills used in listening and reading, yet it differs considerably from these tasks. It is similar to listening in that the rate of

transmission is controlled by the sender. A caption is also fast-fading (Hockett, 1963). The actual signal is present for only a few seconds, and a given caption must be linked to previous information that is not physically available. Caption comprehension is similar to reading, of course, in that the channel of communication is visual, and the medium is written language. It is a form of reading, however, in which lookbacks are impossible. In addition, the captions compete with the accompanying video for attention.

The main goal in captioning is to present captions that are easily understood yet accurately preserve as much of the information in the audio as possible. Under current procedures, captions are simplified with respect to the audio for two primary reasons: (a) Whereas audio information may be processed simultaneously with the video (because these sources are in different channels), captions share the limited-capacity visual channel with other information; and (b) the target audience is assumed to have low-level reading skills (Conrad, 1979; Quigley & King, 1975). Captions are syntactically and lexically simpler than the original audio; idiomatic expressions and metaphors are replaced with concrete language. In effect (although not by design), captioners are following strategies which increase the "readability" of a captioned text as measured by standard readability formulas (e.g., Bormuth, 1966; Dale & Chall, 1948; Flesch, 1948). In addition, portions of the audio are not captioned at all.

The effects of these simplifications on comprehension, retention, interest, and enjoyment are largely unknown. Preliminary analyses of several captioned programs indicate that many current procedures may have negative effects in these areas. Consider the following example, in which the audio portion of a news broadcast is compared to the captions that actually appeared on the screen.

> Original--(Reporter): The energy proposal almost went down the drain today here in the House of Representatives. The battle was over lumping the natural-gas bill with less controversial energy bills. Some senators wanted to split off natural-gas, hoping then to kill it.
>
> (Senator): "I could go on all day reciting the prestigious national organizations that are in all-out opposition to this legislation. And yet, we're being told as a proud legislative body that we should swallow this whole indigestible mass with one single up-or-down vote."
>
> Captioned version--(Reporter): The energy program almost ended today in the House of Representatives. The argument was about linking the natural-gas bill to less controversial energy bills.
>
> (Senator): "I could continue listing the honored national organizations that completely oppose this legislation. But we are being told that we should accept this whole legislation by voting on it once."[7]

The captioned text differs from the original in several ways. It has been shortened by about a third through deletion; individual sentences are shorter and syntactically less complex. Simple lexical items

have been substituted for more difficult ones (e.g., "honored" instead of "prestigious"). Idiomatic expressions have been replaced by non-idiomatic ones (e.g., "ended" replaced the idiom "down the drain").

These alterations, which are typical, affect both the ease with which individual sentences can be decoded and the coherence of the story. Some alterations change the meaning of the text; this is especially true in cases of idiom replacement. For example, saying that "the energy program almost went down the drain" does not mean "the energy program almost ended." The latter implies that an existing energy program was about to end, which is untrue; the former correctly suggests that the bill which would have created the energy program was almost defeated. Similarly, idiomatic phrases in the quotation have been replaced with neutral phrases that fail to convey the speaker's attitude and manner.

Information that may be critical to understanding the story has been deleted entirely. By eliminating the sentence about splitting off natural gas in order to kill it, the explanation for the senator's argument has been lost. The phrase "by voting on it once" in the final sentence of the captioned text is not a rephrase of this missing information; furthermore, it is ambiguous.

The result of these alterations is a text that may fail to convey the story accurately, despite the fact that individual sentences are indeed simpler than the originals. The reader is left with a series of sentences that may be difficult to integrate into a meaningful

interpretation; thus, readability may not have been enhanced. Research is needed to determine which factors contribute to the comprehensibility of captions and to assess how well current captioning practices fit the language-processing abilities of the target audience.

Factors Affecting Caption Comprehension

Display Conventions

In addition to deciding what information to include, the captioner must determine how it is to be displayed. The obvious strategies are not necessarily the most effective. Consider, to take just one example, the question of how to divide a sentence which is to be captioned.

Captions are usually limited to one or two lines so as not to interfere with the video. Because each caption appears on the screen for only a few seconds, the breaks between lines on a single display and between successive displays have more importance than they do in written text, where reading occurs at an individualized pace and look-backs are possible. In order to comprehend the captions, the reader must link the information currently being presented with information presented earlier. Assume, for example, that the text to be captioned is (1). This sentence is too long to be presented on a single line.

(1) President Carter believes that the meeting with Reagan succeeded.

Captioners currently employ two strategies in dividing such sentences into lines. Either the lines are divided so as to equalize the number of words per line (e.g., 2).

(2) President carter believes that the
meeting with Reagan succeeded.

or a break is made at a major clause boundary (e.g., 3).

(3) President Carter believes
that the meeting with Reagan succeeded.

It might be preferable, however, to divide the sentence in a way that takes into account some recent research which shows that listeners and readers use their language and general knowledge in a predictive fashion (Bobrow & Brown, 1975; Marslen-Wilson & Welsh, 1978; Norman & Bobrow, 1975). Thus, the line division in (4) may be preferable because it permits the reader to use the word <u>that</u> in a predictive fashion.

(4) President Carter believes that
the meeting with Reagan succeeded.

By the end of the first line, the reader knows that a complement clause will follow; this knowledge produces a strong expectation that facilitates subsequent processing even if the first line disappears before the complement clause is encountered. Placing the complementizer at the beginning of the second line (as in 3) requires the reader to process back to the previous line in order to comprehend the sentence. This will be especially difficult if that line was on the previous display.

<u>Cues to Sentence Structure</u>

In the service of shortening and simplifying captions, numerous cues to the structure of individual sentences are eliminated. These often

have the effect of making the sentences more difficult to decode. For example, in (5), deleting the complementizer that still leaves a grammatical sentence. Similarly, in (6) the relative pronoun whom can be deleted.

(5) John believes [that] Harry is leaving.

(6) The man [whom] Sidney knows is a thief.

The captioned text may be shortened by deleting such cues. In the reduced version, however, the reader must recover information that otherwise would be explicitly stated. Thus, deletions result in sentences which are shorter but which may be more difficult to comprehend. Similarly, in (7), a causal relationship between two clauses is signalled by the word because. In (8), this information is only implicitly conveyed in two sentences.

(7) Carter met with Congressional leaders to discuss his proposed energy tax package because he desperately needs their support.

(8) Carter met with Congressional leaders to discuss his proposed energy tax package. He desperately needs their support.

It follows from these considerations that caption readability may be improved by adding surface cues that did not appear in the original audio. For example, in (9), the reader must infer that the energy bill almost died because of the battle mentioned in the second sentence. Explicitly stating this information (as in 10) may facilitate comprehension.

(9) The energy proposal almost died today in the House. The battle was over . . .

(10) The energy proposal almost died today in the House because of a battle over . . .

In other cases, deletions may have a positive effect.

(11) Carla went to see her parents, who are living in Chicago.

(12) Carla went to see her parents. They are living in Chicago.

Sentence (11) consists of a complex construction in which a restrictive relative clause is signalled by the word who. In (12), this has been broken into two syntactically simple sentences but without requiring an additional inference. Thus, (12) may be easier to process.

Global Structure Cues

The above examples show that comprehension of a sentence may be slowed or disrupted when cues to its internal structure are deleted. A similar problem exists at the next higher level of analysis, the global structure of the text. Several recent theories describe the internal structure of narrative and expository texts (e.g., Bruce & Newman, 1978; Halliday & Hasan, 1976; Kintsch & van Dijk, 1978; Mandler & Johnson, 1977; Rumelhart, 1975; Stein & Glenn, 1979). Cues to the internal structure of a captioned text are often deleted in the effort to shorten and simplify. This is seen in the following example:

> Captioned version--I first came to the Middle East 11 years ago after the six-day war. Many Israelis thought it would be the last war. This program is not about armies or diplomats. It is about two families caught in the Middle

> East conflict. The program was filmed 3 years ago. But the same feelings remain today as they have for the last 30 years. The feelings remain even though there was hope from the historic visit to Israel by Egypt's President Anwar Sadat. This program was the idea of an Israeli TV producer and a Harvard professor.[8]

The most striking characteristic of the text is that the sentences do not appear to follow from one another. The topic shifts frequently; a topic introduced early in the text (e.g., information about the program) is abruptly reintroduced after much intervening text. The passage does not read as a well-constructed story, but rather as a series of short, disjointed sentences. This may make comprehension more difficult (see Section 4 by Bruce & Rubin).

The story lacks structure because information that provided structure in the original has been deleted. For example, the sentence "This program is not about armies or diplomats" entails a sudden change of topic. In the original, however, it was introduced by a transition sentence, "What you are about to see is the first in a series about the Middle East." This sentence informs the reader about the structure of the story and prepares him/her for elaborative information. Similarly, the sentence "This program was the idea of an Israeli TV producer and a Harvard professor" also involves an abrupt change of topic. Again, however, in the original it was introduced by a transition sentence, "Before we begin this series, a word about how it all came to be and about some of the unusual people who were involved in it."

These examples suggest that introductory material and explicit transitions may facilitate the comprehension process even though they lengthen the text.

Lexical Choice

It is important that captioners use words and expressions that are within the vocabularies of the target audience. Unfortunately, precise information of this kind is not known. A number of studies have shown that hearing-impaired persons typically exhibit low-level reading skills (for review, see Conrad, 1979). On this basis, captioners delete or replace difficult and unusual words. In addition, idiomatic expressions, metaphors, and other fanciful language are deleted or replaced.

This practice presents us with three questions: First, do the low-level reading skills of many hearing-impaired persons reflect a deficit at the vocabulary level, or are they attributable to other factors? Second, what are the capabilities of hearing-impaired persons for understanding non-literal language? Third, what are the consequences of altering the vocabulary items or the nonliteral language of a text?

Expressions are frequently replaced with concrete language which does not preserve the meaning of the original. In one story, the idiom "as time ran out" was replaced with "near the end," which fails to convey the idea of a deadline. In another story, "worked my way through college" was replaced with "worked so I could go to college," which again does

not capture the original meaning. In addition, these expressions are less interesting than the originals; replacement produces a homogeneous text that lacks any style and may be boring to read.

It is also important to evaluate the effects of idiomatic expressions on the comprehension of the longer text in which they are embedded. If idioms are more difficult, they will require time and processing resources which would otherwise be used in analysis of the continuing text. Hence, a "ripple" effect may occur in which overall comprehension of the extended text suffers as well. Note that the opposite could hold instead. Idioms might disrupt processing at a local level but, by adding to the liveliness or character of the text, might at the same time enhance its interest value for the reader. This could lead to increased attentiveness, which would facilitate comprehension of the larger text. Furthermore, the larger text can provide additional information which reveals the meaning of the idiom or permits the viewer to infer it. This issue can only be resolved through direct comparisons of idiom comprehension in limited and extended contexts.

Metaphors may have an even greater effect on the comprehension of longer texts. While a metaphor may increase the difficulty of the sentence in which it occurs, it may facilitate the comprehension of subsequent information by providing a framework for additional detail. This can only be observed if the metaphor is embedded in a longer text.

Genres

Any development of general captioning techniques must be done with the realization that various types of programs may require different guidelines. The experience of WGBH captioners and preliminary analyses of several captioned programs indicate that all of the factors discussed above interact with the genre of the show. For example, for Hawthorne's Scarlet Letter, a prime consideration was to preserve as much as possible of the original language--including most of its metaphors, idioms, and dialectal idiosyncrasies. In contrast, idioms and metaphors are systematically deleted for news captions.

Recent research in text comprehension has focused on the effect of the perceived genre of a text (Brewer, 1980) on readers' expectations and understanding. Green (1979) demonstrated that a newspaper story presented as a narrative is perceived as disconnected and incomprehensible; Adams and Bruce (1981) discuss the knowledge about fables which is crucial to arriving at the "standard" interpretation of Aesop's fables. These genre distinctions have implications for captioning.

On the most general level, adult programs which have been captioned fall into three categories--dramas, documentaries, and news. Although the category of dramatic presentations is hardest to describe in general terms, some captioning issues appear to arise most often within this group. For example, because most dramas have a large amount of dialogue, placing captions so the speaker is identifiable is an issue of prime

concern; off-camera speakers present a difficult problem. In the Miracle Worker, captioners solved the problem of representing the voice of Annie Sullivan's dead brother's ghost by using a smaller font, placing the captions above Annie Sullivan's head and preceding the spoken words with Jemmie's voice. Any guidelines for display conventions will have to be flexible enough to allow for such genre-specific requirements.

Documentaries usually focus either on a person (e.g., Cat, about a female boxer) or a situation (Arabs and Israelis); in either case, it is important to identify early in the show what the topic and the scope of the discussion will be. The original version of the captioned passage from Arabs and Israelis quoted above performed this function by identifying the families from whose perspectives the conflict will be viewed, and indicating when and where the program was filmed. Much of this material provided crucial organizing information. Our preliminary observations indicate that such cues to global structure assume a central role in documentaries.

News programs may seem like the most uniform genre. However, even within a single news show there may be examples of subgenres which require different captioning styles. We will contrast two of these here. The first is the "quickie" story which occurs after the major news is communicated. The anchor person reads the story in 10-15 seconds; often the last sentence is a humorous punch line. The second

is the major news story, a 1- to 2-minute report, often by a correspondent in another city, often complete with short quotations and interviews and some changes in the accompanying video.

The quickie news story presents particular problems for captioners because of its brevity. The density of facts squeezed into such a short time limits the modifications which can be made to the text. Consider the following story:

> Original--James Earl Ray was married today in the visitors' lounge of the Tennessee prison where he is serving a 99-year sentence for the assassination of Dr. Martin Luthur King, Jr. His bride is a 32-year old freelance courtroom artist named Anna Sandhu. Immediately after the ceremony Ray was sent back to his cell and his wife went home. Said Mrs. Ray, "I would like to live life normally."[9]

The story has two main goals: (a) to communicate the facts of Ray's marriage and (b) to deliver a humorous, memorable punch at the end. Reducing the number of words in this story will necessarily involve deleting some concepts which might be crucial to the total message, sacrificing one of these two goals. In the captioned version all of the facts were retained and the punch line deleted. A different decision might have been to sacrifice the details to keep the humor at the end.

Contrast this brief news item with the energy story discussed in the Issues section. For the news item there is only one speaker; for the energy story there are two. When the video changes from the reporter

to the senator, the senator's words must appear on the screen at the same time as his face. The captioner responded to this constraint by deleting the sentence ("Some senators . . .") which had explained the senator's statement. An alternate strategy would have been to combine the first three sentences into two, attempting to retain both the introduction of topic (accomplished in the first and second original sentences) and the explanation of the senator's comments (accomplished by the third sentence).

Summary

The clear point emerging from these preliminary analyses is that producing easily comprehended captions is not simply a matter of making them lexically and syntactically simple. A captioned program presents a text with an internal structure dictated by the script. Local changes in the text may do violence to its overall coherence. This may actually decrease the level of comprehension (or perhaps fail to increase comprehension in the manner intended). Thus: Simple sentences do not necessarily produce simple texts. An individual word, phrase, or sentence may be easily comprehended, while the story in which it is embedded is not.

Discussion

The discussion began with several participants remarking on the role of subtitles in watching foreign films as well as on television in other, non-English-speaking countries.

Selden: Some learning can take place just by observing the video portion of foreign films and matching it with the text.

Seidenberg: Viewers would appear to have some kind of "metaknowledge" of how captions typically relate to the pictures being shown. Such knowledge may mean that the amount of redundancy now in evidence could be reduced. Of course there remains the question of the composition of the target population and whether or not such a population has these skills. The net effect of deletions seems to be to force the perceiver to generate information which was in the original text. That is, by shortening the text, very little is really to be gained if the reader then has to generate more inferences to make sense of the text.

Kantor: Perhaps work needs to be done to determine genre differences between texts. There is obviously a difference between the flowing information of the ABC Evening News and a drama such as The Scarlet Letter. The captioners of the latter broadcast seemed to have little trouble keeping the captions verbatim representations of the audio, principally because the dialogue was brief.

Seidenberg: Captioners are clearly aware of such genre differences.

Rubin: In fact, there was actually a policy decision at the Caption Center to change as little of the language of The Scarlet Letter as possible because of the important effect of style in such a work. (Incidentally, they also captioned a production of The Miracle Worker, resulting in some rather circular effects.) The hardest material to delete occurs in so-called "Peoria stories" (e.g., "There was a fire

this afternoon in Peoria, Illinois, and x number of people died . . .") since they contain little if any deletable information and may take no more than 15 seconds to report. The captioning simply is not able to keep up with the audio.

Seidenberg: And of course, there is the political decision never to eliminate entirely any one news item.

Griffin: Some research has been conducted on little children watching television programs that were beyond their level of comprehension. The initial episode of Superman, for example, in which the baby is launched in a rocket to escape the destruction of the planet Krypton, is extremely confusing. A roomful of young children watching that kind of program reportedly talked throughout it, asking each other questions and helping each other understand. When some caption readers watch television, they make it a practice to form groups, making the viewing a social event. It may be that social collaboration among caption viewers can do some of the work in trying to understand the program, as was the case with the children viewers. While such interaction cannot be assumed, it may be worth investigating how frequent it is and how it affects caption viewing.

Kantor: Since cable television may someday enable individual caption viewers to choose the level and rate of captions by providing one channel for verbatim captioning, one for very simplified captions, and so on, future policy decisions may not have to deal with the problem of "target audience," as is currently the case.

Seidenberg: However, it is doubtful that the commercial networks will adapt programming to suit a minority of viewers who rely on captions. It is really something of a miracle that television is going to gain in importance among this group, but there is a real problem with the way in which captioning is being implemented. There is also a problem with the resolution of the new decoded captions. Apparently the quality of resolution is very poor and it is difficult to read the captions.

8. Final Discussion

The final session of the Conference took the form of an open discussion, which centered around what were felt to be the two main issues at hand: (a) What are the legitimate arguments against the use of readability formulas for certain purposes? and (b) What are sound alternatives to readability formulas, and how can they be supported by empirical evidence?

Arguments Against Some Uses of Readability Formulas

Readability formulas have been immensely popular with publishers and some educators, despite their limitations. Given that there are some problems with the formulas as they now exist, the opponents of the formulas have to deal with why they remain so popular and how to render them less so.

Green: One possible way of demonstrating the problems with formulas is through field testing; field tests might show that material which scores poorly on readability formulas is nevertheless readable.

Selden: Such studies have been done in connection with tax forms, but one cannot immediately generalize to children results that are obtained with adults, or vice versa. Studies specifically designed to challenge the use of formulas on children's texts by showing that some texts which score well do not "work" as well as texts which score poorly, probably should be done. A variety of tasks could be used to measure how well children understand a text. It is imperative that the specific task be defined before any test of readability is devised, and the task must be an appropriate one. Cloze tests, for example, are not particularly natural tasks for children.

Kantor: Individual studies might be aimed at showing that one particular effect of the application of readability formulas is bad, e.g., sentence splitting.

Selden: But it is clear, however, that formulas are not intended to be used as guides for writing, so this type of criticism would only apply to improper uses of the formulas.

Bruce: The burden of proof may well lie with the proponents of the formulas. To what extent are the formulas better predictors than the average school teacher or librarian?

Charrow: The formulas have been revised a number of times, presumably to make them more and more accurate.

Selden: But validation studies have been unbelievably weak, validating one formula against another. Furthermore, the results have been generalized far beyond the initial texts.

Griffin: Basic to the use of readability formulas has been the notion that statistics, numbers, carry more truth than do other kinds of facts. The use of reading levels relies on a kind of generalization that if someone can read at a given level, e.g., 6.2, then he or she can read anything that measures below that level. This generalization is unwarranted because the formulas are very text specific.

Bruce: The effect of cultural backgrounds is further evidence that the generalization of reading levels is not quite right. On such tests as the BEAT and BITCH, for example, blacks outperform whites. This points to factors such as culture which influence the "readability" of a text, but which are necessarily ignored by the formulas.

Seidenberg: The pressures for the use of readily applicable formulas appear to supercede all these considerations. It is obvious that, given the needs of people who are under state mandates to use formulas, formulas are simply a fact of life. It is another issue, however, to determine whether techniques are necessary which permit us to predict ahead of time how readable a text is. Good writing, apparently the result of talent, is clearly beyond the predictive realms of the formulas. Likewise, the ethics of whether or not an adaptor has the right to tamper with the content of a text--deleting information, for example--has not been addressed by the proponents of formulas.

Davison: It is probably a mistake to mix the politics and economics of readability formulas with the legitimate logical arguments against them.

Charrow: The politics and economics are in their favor, particularly since the powers that be (e.g., legislators) really believe that formulas are an accurate gauge of comprehension.

Davison: The most we can hope to accomplish is to spell out a systematic attack based on the logical arguments against readability formulas in their present state. With such a list in hand, we could then go about the business of convincing others that there is a strong case to be made against the indiscriminate use of formulas. That, at least, would be a beginning.

The major logical arguments against readability formulas are as follows:

1. Morgan (Linguistics, University of Illinois): The validation studies have been done are both weak and circular in nature. What little data there is tends to validate one formula against another.

2. Morgan: The mistake is made of particularizing from the average to a single sentence--i.e., from mass statistics to one text and one reader. It is assumed that since mass statistics show a general tendency, one can particularize to one passage for one child. This is a non-sequitur.

3. Bruce: Related to 2 is the questionable practice of generalizing or extrapolating from the original passage upon which the readability formula is based to any passage. It is poor statistical method to extrapolate to another population from that of the original study.

4. Morgan: The parameters which are measured by readability formulas are incomplete. There are numerous factors which go unmeasured by the formulas but which contribute to readability. These include specific syntactic characteristics, pragmatic considerations (e.g., inferences), cultural aspects, how enjoyable a text is to read, and so forth.

5. Morgan: It is probably unrealistic to expect any formula to deal with all of the factors which affect readability. A simple mechanical measure will be limited to certain characteristics of a text. Therefore, no readability formula should be regarded as adequate for measuring readability.

6. Selden: Related to (5) is the caveat that if a formula is designed to measure a given task, A, on a given group, B, it should only be used for that. Because the formulas are so widely believed in, there is a tendency to misuse them. Included in the misapplication of readability formulas is their use as guides to writing or adapting, acknowledged as a misapplication by the designers of the formulas.

Much of the discussion centered around 4 and the notion that certain parameters could not be included in formulas, at least not without making them into cumbersome and hopelessly complicated mechanisms. Whatever gains would be made by attempting to include such factors would be vitiated by the difficulties which the new formulas would create. (They would cease to be quick and easy.)

Davison: What becomes increasingly clear, however, is just how imprecise the formulas are at present. The notion of grade levels dates back to the era of the McGuffy Readers. The levels are now taken for granted to be meaningful, and readability formulas measure texts against them. Just how absurd this notion is should be evident when the formulas are used to determine the "grade level" of a tax form intended for adults who are required to perform a specific task (unrelated to school texts in both form and purpose). In fact, it may on occasion be more difficult to read at a lower level than one is used to, since much of the information from which to draw inferences would be lacking and the reader would need to do more work in order to extract the same amount of information.

A. Zwicky, Columbus, Ohio: While it is clearly not advisable to give children passages which are much too difficult for them, some challenge, at least, might be beneficial and help improve the child's reading ability. Furthermore, despite improvements in texts from the standpoint of the formulas, there is little evidence that reading scores have been improving over the years. Reading scores may, in fact, be deteriorating, though this is difficult to determine due to changes in language, population, and culture over the years.

Alternatives to Readability Formulas

Bruce: For all of the reasons mentioned above, many people do feel a need for some standard of measuring how readable texts are. One

obvious alternative to formulas is to field-test materials. By giving a text to people and actually seeing whether the text causes comprehension problems, we can achieve results which will probably be as good as, if not better than, the information provided by the standard readability formulas.

Davison: The formulas are now being used to handle such texts as contracts and tax forms--designed, of course, for adult readers--and it is imperative that the content of these texts be easily understood by the readers. Since in studying these we would be dealing with a relatively small corpus of texts (unlike the situation with basal readers or trade books for children), with a relatively closed set of instructions for the reader, field testing seems both appropriate and feasible.

Selden: Publishers currently spend large sums of money to have programs analyzed by readability formulas. That same money could be put to better use by having their new texts field tested.

Tierney: Currently, materials designed for the classroom must be approved of by teachers, and this itself is a kind of field test. Field testing was in fact done in the Natural Assessment program by Petroski (Personal communication) at the University of Pittsburgh in conjunction with school librarians. Classes of 9- and 13-year-olds were allowed to make comments on readings and asked to eliminate passages which they disliked. The criteria were how enjoyable a text was and how well suited it was to various purposes.

Morgan: Of all the skills taught to children, only reading is subject to the supposedly objective method of the readability formulas.

Griffin: This is in part due to the hierarchical nature of such areas as mathematics, where the basic skills of addition and subtraction are considered to be essential to the acquisition of more advanced skills such as multiplication and division.

Selden: But there is also a sense that readability formulas have gotten out of hand, since they were originally no more than an attempt to control the vocabulary in basal readers, e.g., Thorndike's word list, for use in vocabulary recognition skills.

Charrow: Parallel to the development of math skills, there has been an attempt to build a hierarchy into the reading process.

Tierney: Teachers are told that the formulas are used as a yardstick which, on an analogy with a shoe store, will give each child a custom fit. The application of Harris' five-finger rule--i.e., more than five words per page which cause the child problems means that the text is inappropriate--is part of this idea of custom fit.

Kantor: By field testing, it would be possible to determine just what syntactic structures cause problems for the target readers.

Osborn: This would certainly have a high degree of validity and would relate to what the readers actually were able to handle.

Tierney: But again, we are faced with the pragmatic consideration that this might not be the "quick and dirty" method which people seem to want.

Griffin: Still, it ought to be possible to add on some of these "higher-level" skills to the already existing hierarchy. Whether or not the hierarchy idea itself is defendable has been the subject of much controversy.

Tierney: Publishers are being forced to homogenize texts--i.e., level out the difficulties and make all the passages conform to an average.

Davison: Such a trend seems both pernicious and unnecessary.

Selden: Regulating language in this way provides such groups as insurance companies with numbers which "prove" that a form conforms to certain averages. This is an easy way to reply to challenges which might arise that a form is, for instance, too difficult). Similarly, school boards can point to the formulas when challenged that their reading programs have failed. Formulas are called upon despite the discrepancies which exist between the scores according to the various formulas. The formulas look objective, people think that they are measuring something of importance, and they are quick and easy to apply. Thus, the formulas remain incredibly attractive to many, even when misapplied to rewrite texts (e.g., the case of the IRS splitting long sentences even when this resulted in instructions which were less clear than the original). Communicability seems to be less of a priority than meeting some standard of "readability."

Tierney: One concrete step which needs to be taken is the dissemination of such research as the technical report on adaptation written by the Text Analysis Group at the Center for the Study of Reading (Davison et al., 1980).

Osborn: That particular report is quite lengthy, and it would be possible to extract some of the suggestions it contains and send the short report to publishers and others in an attempt to "raise their consciousness" and maybe eventually meet with some of them and discuss the issues.

Bruce: There is also the need for studies which compare the results of readability formulas with those of actual field testing. It would be possible, for example, to devise an experiment which compared the reading level of certain trade books for children with actual field tests.

Selden: It would first be necessary to define the kind of task which is being investigated. This would determine how appropriate the text is for a given task only.

Osborn: Basal readers, however, are felt by many at this conference to be the primary target of our attack.

Charrow: In terms of the rewriting of texts, it is very possible to use general guidelines instead of the rigid formulas.

Bruce: For example, avoid certain constructions under certain conditions (not unlike Strunk & White).

Davison: SRA already uses the guidelines of Dawkins's (1975) compendium, though the compendium really doesn't explain why certain things are complex. What seems to be needed are more research-based guidelines.

In conclusion, there appear to be a number of ways to determine the suitability of any given text besides the application of readability formulas. Guidelines can be drawn up which list principles that should be applied when writing texts for given purposes and readers. Field testing can be done to determine the level of difficulty of an existing text, its appropriateness for well-defined tasks and audiences, how enjoyable it is, and even what could be done to improve it. Teachers should be consulted; their insight is valuable and stems from actual use of reading materials in classrooms. Children themselves are able to provide useful judgments. Syntactic and other parameters could be incorporated into the already existing hierarchies, thus making it possible to better select appropriate texts. In the meantime, our task is to make available to others the logical arguments which we have presented here against the use of readability formulas. Workshops are needed for both teachers and publishers.

Reference Notes

1. Fagan, W. T. The relationship between reading difficulty and the number and type of sentence transformations. Paper presented at the 16th Annual Meeting of the International Reading Association, Atlantic City, N.J., April 1971.

2. Schmidt, E. L. What makes reading difficult: The complexity of structures. Paper presented at the 27th Annual Meeting of the National Reading Conference, New Orleans, L.A., December 1977.

3. Selden, R. Development of a frequency-based measure of syntactic difficulty for estimating readability. Paper presented at the 22nd Annual Meeting of the International Reading Association, Miami Beach, Florida, May 1977.

4. Etter, D. The impact of management and turn-allocation activities on student performance (Sociolinguistic Working Paper No. 65). Austin, Tex.: Southwest Educational Development Laboratory, September 1979.

5. Mellinkoff, D. Some misconceptions about the clear writing of documents. Paper presented at the Conference of Experts in Clear Legal Drafting, National Center for Administrative Justice, Washington, D.C., June 2, 1978.

6. O'Bryan, K. G. Eye movement research report on captioning television programs for the deaf. Unpublished paper, Ontario Institute for Studies in Education, 1976.

References

Adams, M. J., & Bruce, B. Background knowledge and reading comprehension. In J. Langer & T. Smith Burke (Eds.), Author meets reader/A psycho-linguistic perspective: Bridging the gap from theory to practice. Newark, Del.: International Reading Association, 1981.

Bendick, J., & Gallant, R. Elementary science 6. Lexington, Mass.: Ginn & Co., 1980.

Bickman, L., Roth, T., Szoc, R., Normoyle, J., Shutterly, J., & Wallace, W. D. An evaluation of captioned television for the deaf: Final Report. Evanston, Ill.: Westinghouse Evaluation Institute, 1979.

Bobrow, R., & Brown, J. S. Systematic understanding: Synthesis, analysis, and contingent knowledge in specialized understanding systems. In D. G. Bobrow & A. Collins (Eds.), Representation and understanding: Studies in cognitive science. New York: Academic Press, 1975.

Bormuth, J. R. Readability: A new approach. Reading Research Quarterly, 1966, 1, 79-132.

Botel, M., & Granowsky, A. A formula for measuring syntactic complexity: A directional approach. Elementary English, 1972, 49, 513-516.

Brewer, W. F. Literary theory, rhetoric, stylistics: Implications for psychology. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), Theoretical issues in reading comprehension. Hillsdale, N.J.: Erlbaum, 1980.

Brown, L. New device calls up printed matter on TV. The New York Times, February 5, 1980, p. C24.

Bruce, B. C., & Newman, D. Interacting plans. Cognitive Science, 1978, 2, 195-233.

Clymer, T., Parr, B., Gates, D., & Robison, E. G. Ginn 720 Series. A duck is a duck. Lexington, Mass.: Ginn, 1976.

Clymer, T., Wong, O., & Benedict, V. J. Series 720, Level 6, Unit 3, One to grow on. Lexington, Mass.: Ginn, 1976. Pp. 89-91.

Conrad, R. The deaf school child. London: Macmillan, 1979.

Dale, E., & Chall, J. S. A formula for predicting readability. Educational Research Bulletin, 1948, 27, 37-54.

Dale, E., & O'Rourke, J. The living word vocabulary. Elgin, Ill.: Field Enterprises Educational Corporation, 1976.

Davila, R. R. Effects of changes in visual information patterns using a captioned film and specially adapted still pictures. Doctoral dissertation, Syracuse University, 1972.

Davison, A., Kantor, R. N., Hannah, J., Hermon, G., Lutz, R., & Salzillo, R. Limitations of readability formulas in guiding adaptations of texts (Tech. Rep. No. 162). Urbana: University of Illinois, Center for the Study of Reading, March 1980. (ERIC Document Reproduction Service No. ED 184 090)

Dawkins, J. Syntax and readability. Newark, Del.: International Reading Association, 1975.

Document Design Project. Simplifying documents: A three-day workshop. Washington, D.C.: American Institutes for Research, 1979.

Dolch, E. W. Vocabulary burden. Journal of Educational Research, 1928, 17, 170-183.

Dolch, E. W. Grading reading difficulty. In E. W. Dolch (Ed.), Problems in reading. Champaign, Ill.: Garrard Press, 1948. Ch. 12.

Education Progress Corp. Clue Magazine, No. 2, 1972, 1979.

Educational Testing Service. Cooperative English test of reading comprehension. Princeton, N.J.

Endicott, A. L. A proposed scale for syntactic complexity. Research and the Teaching of English, 1973, 7, 5-12.

Fereirro, E. What is written in a written sentence: A developmental answer. Journal of Education, 1978, 160, 25-39.

Fischer, D. C. Improvement in the utilization of captioned films for the deaf. Doctoral dissertation, University of Nebraska, 1971.

Flesch, R. A new readability yardstick. Journal of Applied Psychology, 1948, 32, 221-233.

Flesch, R. The art of readable writing. New York: Harper, 1949.

Flesch, R. How to test readability. New York: Harper, 1951.

Flesch, R. How to make sense. New York: Harper, 1954.

Fodor, J. A., & Garrett, M. Some syntactic determinants of sentential complexity. Perception & Psychophysics, 1967, 2, 289-296.

Freeman, C., Larkin, D., Humphrey, F., & Yanofsky, N. Proposal to the National Institute of Education from the Center for Applied Linguistics. Arlington, Va.: 1978.

Freire, P. Cultural action for freedom. Harvard Educational Review, Monograph Series (No. 1). Cambridge, Mass.: Harvard Educational Review, 1970.

Fry, E. B. A readability formula that saves time. The Journal of Reading, 1968, 11, 513-516; 575-578.

Gates, R. R. The differential effectiveness of various modes of presenting verbal information to deaf students through modified television formats. Doctoral dissertation, University of Pittsburgh, 1970.

Gourley, J. This basal is easy to read--or is it? The Reading Teacher, 1978, 32, 174-182.

Gourley, J., & Catlin, J. Children's comprehension of grammatical structures in context. Journal of Psycholinguistic Research, 1978, 7, 419-434.

Granowsky, A., & Botel, M. Background for a new syntactic complexity formula. The Reading Teacher, 1974, 28, 31-35.

Green, G. Organization, goals, and comprehensibility in narratives: Newswriting, a case study (Tech. Rep. No. 132). Urbana: University of Illinois, Center for the Study of Reading, July 1979. (ERIC Document Reproduction Service No. ED 174 949)

Gunning, R. The technique of clear writing. New York: McGraw-Hill, 1952.

Gunning, R. How to take the fog out of writing. Chicago: Follett, 1964.

Halliday, M. A. K., & Hasan, R. Cohesion in English. London: Longman, 1976.

Harris, A. J., & Jacobson, M. D. Revised Harris-Jacobson readability formulas. Bethesda, Md.: College Reading Association, 1974.

Hess, R. D., & Takanishi, R. The relationship of teacher behavior and school characteristics to student engagement (Tech. Rep. No. 42). Stanford, Cal.: Stanford Center for Research and Development in Teaching, 1974.

Hockett, C. The problem of universals in language. In J. Greenberg (Ed.), Universals of language. Cambridge, Mass.: MIT Press, 1963.

Hunt, K. Grammatical structures written at three grade levels. Champaign, Ill.: National Council of Teachers of English, 1965.

Hutchins, P. Don't forget the bacon. New York: Greenwillow Books, 1976.

Jacobson, M. D., Kirkland, C. E., & Selden, R. An examination of the McCall-Crabbs standard test lessons in reading. Journal of Reading, 1978, 22, 224-230.

Kintsch, W. Memory and cognition. New York: Wiley, 1977.

Kintsch, W., & van Dijk, T. A. Toward a model of text comprehension and production. Psychological Review, 1978, 85, 363-394.

Kintsch, W., & Vipond, D. Reading comprehension and readability in educational practice and psychological theory. In L. G. Nillson (Ed.), Memory processes. Hillsdale, N.J.: Erlbaum, 1979.

Klare, G. R. The measurement of readability. Ames: Iowa State University Press, 1963.

Klare, G. R. Assessing readability. Reading Research Quarterly, 1974-75, 10, 62-102.

Kozol, J. A new look at the literacy campaign in Cuba. Harvard Educational Review, 1978, 48, 341-377.

Lewerenz, A. S. Measurement of the difficulty of reading materials. Educational Research Bulletin, 1929, 8, 11-16.

Lively, B. A., & Pressey, S. L. A method for measuring the "vocabulary burden" of textbooks. Educational Administration and Supervision, 1923, 9, 389-398.

Luria, A. R. The development of writing in the child. Soviet Psychology, 1977-78, 16(2), 46-64.

Mandler, J. M., & Johnson, N. S. Remembrance of things parsed: Story structure and recall. Cognitive Psychology, 1977, 9, 111-151.

Markova, A. K. The teaching and mastery of language. White Plains, N.Y.: Sharpe, 1979.

Marslen-Wilson, W. D., & Welsh, A. Processing interactions and lexical access during word recognition in continuous speech. Cognitive Psychology, 1978, 10, 29-63.

McCall, W. A., & Crabbs, L. M. Standard test lessons in reading. Teachers College Record, 1925, 27. (a)

McCall, W. A., & Crabbs, L. M. Standard test lessons in reading, Books 2, 3, 4, and 5. New York: Teachers College Press, Columbia University, 1925. (b)

McCall, W. A., & Crabbs, L. M. Standard test lessons in reading, books A, B, C, D, and E. New York: Teachers College Press, Columbia University, 1961.

McKenzie, M. The beginnings of literacy. Theory Into Practice, 1977, 16, 315-324.

Meyer, B. J. F. The organization of prose and its effect on recall. Amsterdam: North-Holland, 1975.

Nelson, J. Readability: Some cautions for the content area teacher. Journal of Reading, 1978, 21, 620-625.

Nix, G. W. The effects of synchronized captioning on the assimilation of vocabulary and concepts presented in a film to intermediate level deaf children. Unpublished doctoral dissertation, University of Oregon, 1971.

Norman, D. A., & Bobrow, D. G. On data-limited and resource-limited processes. Cognitive Psychology, 1975, 7, 44-64.

Norwood, M. J. Captioned films for the deaf. Exceptional Children, 1976, 43, 164-166.

Norwood, M. J. Comparison of an interpreted and captioned newscast among deaf high school graduates. Unpublished doctoral dissertation, University of Maryland, 1976.

O'Bryan, K. G. Captioning television programs for the deaf: A report of basic findings. OCEA Report, 1975.

Paley, V. G. Wally's Stories. Cambridge, Mass.: Harvard Press, 1981.

Palmer, W. S. Readability, rhetoric, and the reduction of uncertainty. Journal of Reading, 1974, 17, 552-558.

Parker, D. A. SRA reading laboratory. Secondary Series, 111b, Chicago: Science Research Associates, 1963.

Propp, G. An experimental study on the encoding of verbal information for visual transmission to the hearing impaired learner. Doctoral dissertation, University of Nebraska, 1972.

Quigley, S., & King, C. The language development of deaf children and youth. Urbana: University of Illinois, 1975.

Reddin, E. Syntactical structure and reading comprehension. Reading Teacher, 1970, 23, 467-469.

Richek, M. A. Effect of sentence complexity on the reading comprehension of syntactic structures. Journal of Educational Psychology, 1976, 68, 800-806.

Rumelhart, D. E. Notes on a schema for stories. In D. G. Bobrow & A. Collins (Eds.), Representation and understanding: Studies in cognitive science. New York: Academic Press, 1975.

Schein, J. D., & Delk, M. T., Jr. The deaf population of the United States. Silver Spring, Md.: National Association of the Deaf, 1974.

Selden, R. Sentence structure and readability: The influence of frequency-based expectancies of syntactic structure on the comprehensibility of text. Unpublished doctoral dissertation, University of Virginia, 1977.

Sendelbaugh, J. Television viewing habits of hearing-impaired teenagers in the Chicago metropolitan area. American Annals of the Deaf, 1978, 123, 536-541.

Simpkins, G., Holt, G., & Simpkins, C. Shine. In Bridge Series: A crossculture-reaching program (Booklet 1). Boston: Houghton Mifflin, 1977.

Spache, G. A new readability formula for primary grade reading materials. Elementary School Journal, 1953, 53, 410-413.

Stein, N., & Glenn, C. G. An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), New directions in discourse processing. Norwood, N.J.: Ablex, 1979.

Steinberg, C. S., & Bruce, B. C. Higher level features in children's stories: Rhetorical structure and conflict. The National Reading Conference Yearbook. Clemson, S.C.: National Reading Conference, 1980.

Sticht, T. G. Project realistic: Determination of adult functional literacy skill levels. Reading Research Quarterly, 1972, 7, 424-465.

Sticht, T. G. (Ed.) Reading for working. Alexandria, Va.: Human Resources Research Organization, 1975. Chapter 2.

Thorndike, E. A teacher's word book of the twenty thousand words found most frequently and widely in general reading for children and young people. New York: Teachers College Press, 1920.

Vogel, M., & Washburne, C. An objective method of determining grade placement of children's reading material. Elementary School Journal, 1928, 28, 373-381.

von Glaserfeld, E. The problem of syntactic complexity. Journal of Reading Behavior, 1970-71, 3, 1-14.

Walmsley, S. On the purpose and content of secondary reading programs: An educational and ideological perspective. Curriculum Inquiry, 1981, 2, 73-93.

Winslow, L. A. Formative evaluation of captioned Zoom. Report prepared for WGBH-TV, Boston, 1977.

Wolf, T. A new theory of skilled readings. Unpublished doctoral dissertation, Harvard University, 1976.

Footnotes

[1]The content of this paper does not necessarily reflect official NIE policy.

[2]Credits beyond specific citations are due to Cissy Freeman and Sean Walmsley for both ideas and arguments.

[3]From "Arabs and Israelis: Two Families," a WGBH-TV production, April 13, 1978.

[4]For additional information about captioning, see Section 7, Comprehension of Captioned Television, which was presented at the Conference by Mark Seidenberg and co-authored with B. C. Bruce and Andee Rubin.

[5]The 17 guidelines are listed in Document Design (1979). A revised edition is now in press.

[6]This is part of a preliminary report written by Mark Seidenberg, Bertram C. Bruce, and Andee Rubin at Bolt, Beranek and Newman, Inc., Cambridge, Mass. More reports are forthcoming. Thanks to the Caption Center, WGBH-TV Boston, for making available captioned materials and for discussing those issues with us.

[7]From the Captioned ABC News, October 13, 1978.

[8]From "Arabs and Israelis: Two Families," a WGBH-TV production, April 13, 1978.

[9]From the Captioned ABC News, October 13, 1978.

Table 1

Comparison of Reading Levels

	California's Giants		Disaster in Dayton		Milk		Living Light	
	Fry	Dale-Chall	Fry	Dale-Chall	Fry	Dale-Chall	Fry	Dale-Chall
Reading Level:								
Original	5	5-6	4+	4.9	10	11-12	11	11-12
Adaptation	6	5-6	2	4-	5	7-8	8	7-8

Table 2

Average Number of Words Per Sentence

	Milk	Light	Dayton	Trees
Original	24	19	12.5	12.5
Adapted	13	14	9	12

Table 3

Comparison of Word and Sentence Length

	California's Giants	Disaster in Dayton	Milk	Living Light
Number of Words:				
Original	3,000*	4,725	1,256	1,450
Adaptation	775	827	861	900
Number of Sentences:				
Original	250*	375*	58	75
Adaptation	62	90	63	61

* estimate

Table 4

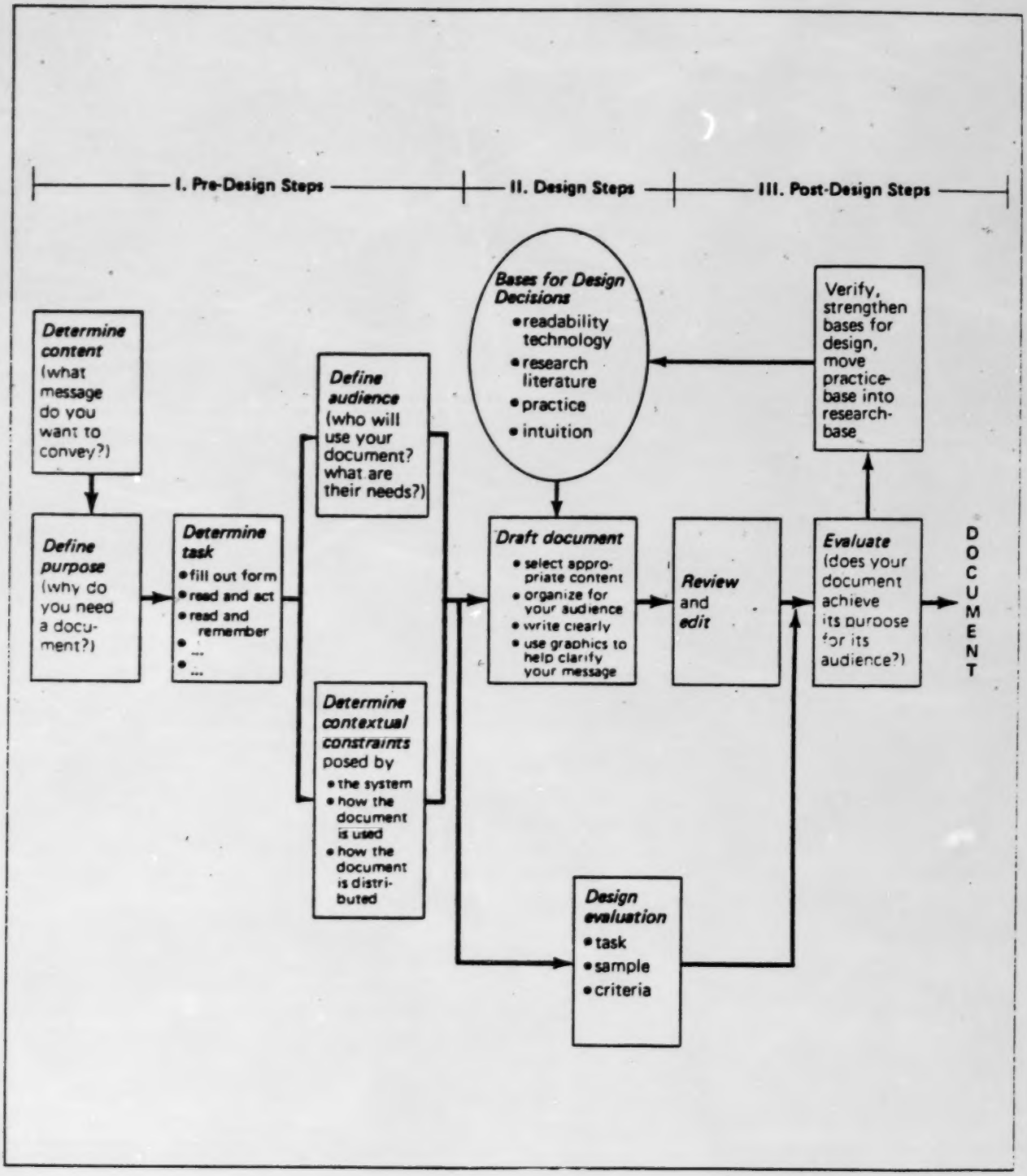
Clause Complexity in LIGHT and MILK

Number of Sentences in:	Number of Clauses Per Sentence					
	1	2	3	4	5	6+
Light						
Original	21	22	23	7	3	-
Adapted	22	20	12	-	-	-
Milk						
Original	20	16	12	2	3	6
Adapted	29	21	15	6	-	-

Note. In counting number of clauses, we have included subject, object and adverbial complement clauses, and relative clauses, including reduced post-nominal modifiers and parentheticals. We have excluded prenominal modifiers, nominalizations, and conjoined noun phrases, even though these might have been derived from some more complex source which involved clauses. Our criterion was essentially whether there is a major constituent break in surface structure. Though we have been arbitrary in a few cases, the measure we give is, in general, consistent.

Figure Caption

Figure 1. The Process Model of Document Design.


I. Pre-Design Steps
II. Design Steps
III. Post-Design Steps
Determine content (what message do you want to convey?)
Define purpose (why do you need a document?)
Determine task
• fill out form
• read and act
• read and remember
• ...
• ...
Define audience (who will use your document? what are their needs?)
Determine contextual constraints posed by
• the system
• how the document is used
• how the document is distributed
Bases for Design Decisions
• readability technology
• research literature
• practice
• intuition
Draft document
• select appropriate content
• organize for your audience
• write clearly
• use graphics to help clarify your message
Review and edit
Design evaluation
• task
• sample
• criteria
Evaluate (does your document achieve its purpose for its audience?)
Verify, strengthen bases for design, move practice-base into research-base
DOCUMENT

CENTER FOR THE STUDY OF READING

READING EDUCATION REPORTS

Adams, M. J., Anderson, R. C., & Durkin, D. Beginning Reading: Theory and Practice (No. 3), November 1977. (ERIC Document Reproduction Service No. ED 151 722, 15p., PC-$2.00, MF-$.91)

Adams, M., & Bruce, B. Background Knowledge and Reading Comprehension (No. 13), January 1980. (ERIC Document Reproduction Service No. ED 181 431, 48p., PC-$3.65, MF-$.91)

Anderson, R. C., & Freebody, P. Vocabulary Knowledge and Reading (No. 11), August 1979. (ERIC Document Reproduction Service No. ED 177 470, 52p., PC-$5.30, MF-$.91)

Anderson, T. H. Another Look at the Self-Questioning Study Technique (No. 6), September 1978. (ERIC Document Reproduction Service No. ED 163 441, 19p., PC-$2.00, MF-$.91)

Anderson, T. H., Armbruster, B. B., & Kantor, R. N. How Clearly Written are Children's Textbooks? Or, Of Bladderworts and Alfa (includes a response by M. Kane, Senior Editor, Ginn and Company) (No. 16), August 1980. (ERIC Document Reproduction Service No. ED 192 275, 63p., PC-$5.30, MF-$.91)

Armbruster, B. B., & Anderson, T. H. Content Area Textbooks (No. 23), July 1981.

Asher, S. R. Sex Differences in Reading Achievement (No. 2), October 1977. (ERIC Document Reproduction Service No. ED 146 567, 30p., PC-$3.65, MF-$.91)

Baker, L. Do I Understand or Do I not Understand: That is the Question (No. 10), July 1979. (ERIC Document Reproduction Service No. ED 174 948, 27p., PC-$3.65, MF-$.91)

Bruce, B. What Makes a Good Story? (No. 5), June 1978. (ERIC Document Reproduction Service No. ED 158 222, 16p., PC-$2.00, MF-$.91)

Bruce, B. A New Point of View on Children's Stories (No. 25), July 1981.

Bruce, B., & Rubin, A. Strategies for Controlling Hypothesis Formation in Reading (No. 22), June 1981.

Collins, A., & Haviland, S. E. Children's Reading Problems (No. 8), June 1979. (ERIC Document Reproduction Service No. ED 172 188, 19p., PC-$2.00, MF-$.91)

Davison, A. Readability--Appraising Text Difficulty (No. 24), July 1981.

Durkin, D. Comprehension Instruction--Where are You? (No. 1), October 1977. (ERIC Document Reproduction Service No. ED 146 566, 14p., PC-$2.00, MF-$.91)

Durkin, D. What is the Value of the New Interest in Reading Comprehension? (No. 19), November 1980. (ERIC Document Reproduction Service No. ED 198 499, 51p., PC-$5.30, MF-$.91)

Durkin, D. Reading Comprehension Instruction in Five Basal Reader Series (No. 26), July 1981.

Jenkins, J. R., & Pany, D. Teaching Reading Comprehension in the Middle Grades (No. 4), January 1978. (ERIC Document Reproduction Service No. ED 151 756, 36p., PC-$3.65, MF-$.91)

Joag-dev, C., & Steffensen, M. S. Studies of the Bicultural Reader: Implications for Teachers and Librarians (No. 12), January 1980. (ERIC Document Reproduction Service No. ED 181 430, 28p., PC-$3.65, MF-$.91)

McCormick, C., & Mason, J. What Happens to Kindergarten Children's Knowledge about Reading after a Summer Vacation? (No. 21), June 1981.

Pearson, P. D., & Kamil, M. L. Basic Processes and Instructional Practices in Teaching Reading (No. 7), December 1978. (ERIC Document Reproduction Service No. ED 165 118, 29p., PC-$3.65, MF-$.91)

Rubin, A. Making Stories, Making Sense (includes a response by T. Raphael and J. LaZansky) (No. 14), January 1980. (ERIC Document Reproduction Service No. ED 181 432, 42p., PC-$3.65, MF-$.91)

Schallert, D. L., & Kleiman, G. M. Some Reasons Why Teachers are Easier to Understand than Textbooks (No. 9), June 1979. (ERIC Document Reproduction Service No. ED 172 189, 17p., PC-$2.00, MF-$.91)

Steinberg, C., & Bruce, B. Higher-Level Features in Children's Stories: Rhetorical Structure and Conflict (No. 18), October 1980. (ERIC Document Reproduction Service No. ED 198 474, 27p., PC-$3.65, MF-$.91)

Taylor, M., & Ortony, A. Figurative Devices in Black Language: Some Socio-Psycholinguistic Observations (No. 20), May 1981.

Tierney, R. J., & LaZansky, J. The Rights and Responsibilities of Readers and Writers: A Contractual Agreement (includes responses by R. N. Kantor and B. B. Armbruster) (No. 15), January 1980. (ERIC Document Reproduction Service No. ED 181 447, 32p., PC-$3.65, MF-$.91)

Tierney, R. J., Mosenthal, J., & Kantor, R. N. Some Classroom Applications of Text Analysis: Toward Improving Text Selection and Use (No. 17), August 1980. (ERIC Document Reproduction Service No. ED 192 251, 43p., PC-$3.65, MF-$.91)

CENTER FOR THE STUDY OF READING

TECHNICAL REPORTS

Adams, M. J. Failures to Comprehend and Levels of Processing in Reading (No. 37), April 1977. (ERIC Document Reproduction Service No. ED 145 410, 51p., PC-$5.30, MF-$.91)

Adams, M. J. Models of Word Recognition (No. 107), October 1978. (ERIC Document Reproduction Service No. ED 163 431, 93p., PC-$6.95, MF-$.91)

Adams, M. J. What Good is Orthographic Redundancy? (No. 192), December 1980. (ERIC Document Reproduction Service No. ED 199 663, 74p., PC-$5.30, MF-$.91)

Adams, M. J., & Collins, A. A Schema-Theoretic View of Reading Comprehension (No. 32), April 1977. (ERIC Document Reproduction Service No. ED 142 971, 49p., PC-$3.65, MF-$.91)

Alessi, S. M., Anderson, T. H., & Biddle, W. B. Hardware and Software Considerations in Computer Based Course Management (No. 4), November 1975. (ERIC Document Reproduction Service No. ED 134 928, 21p., PC-$2.00, MF-$.91)

Alessi, S. M., Anderson, T. H., & Goetz, E. T. An Investigation of Lookbacks During Studying (No. 140), September 1979. (ERIC Document Reproduction Service No. ED 177 494, 40p., PC-$3.65, MF-$.91)

Anderson, R. C. Schema-Directed Processes in Language Comprehension (No. 50), July 1977. (ERIC Document Reproduction Service No. ED 142 977, 33p., PC-$3.65, MF-$.91)

Anderson, R. C., & Freebody, P. Vocabulary Knowledge (No. 136), August 1979. (ERIC Document Reproduction Service No. ED 177 480, 71p., PC-$5.30, MF-$.91)

Anderson, R. C., Goetz, E. T., Pichert, J. W., & Halff, H. M. Two Faces of the Conceptual Peg Hypothesis (No. 6), January 1976. (ERIC Document Reproduction Service No. ED 134 930, 29p., PC-$3.65, MF-$.91)

Anderson, R. C., & Pichert, J. W. Recall of Previously Unrecallable Information Following a Shift in Perspective (No. 41), April 1977. (ERIC Document Reproduction Service No. ED 142 974, 37p., PC-$3.65, MF-$.91)

Anderson, R. C., Pichert, J. W., Goetz, E. T., Schallert, D. L., Stevens, K. C., & Trollip, S. R. Instantiation of General Terms (No. 10), March 1976. (ERIC Document Reproduction Service No. ED 134 933, 30p., PC-$3.65, MF-$.91)

Anderson, R. C., Pichert, J. W., & Shirey, L. L. Effects of the Reader's Schema at Different Points in Time (No. 119), April 1979. (ERIC Document Reproduction Service No. ED 169 523, 36p., PC-$3.65, MF-$.91)

Anderson, R. C., Reynolds, R. E., Schallert, D. L., & Goetz, E. T. Frameworks for Comprehending Discourse (No. 12), July 1976. (ERIC Document Reproduction Service No. ED 134 935, 33p., PC-$3.65, MF-$.91)

Anderson, R. C., Spiro, R. J., & Anderson, M. C. Schemata as Scaffolding for the Representation of Information in Connected Discourse (No. 24), March 1977. (ERIC Document Reproduction Service No. ED 136 236, 18p., PC-$2.00, MF-$.91)

Anderson, R. C., Stevens, K. C., Shifrin, Z., & Osborn, J. Instantiation of Word Meanings in Children (No. 46), May 1977. (ERIC Document Reproduction Service No. ED 142 976, 22p., PC-$2.00, MF-$.91)

Anderson, T. H. Study Skills and Learning Strategies (No. 104), September 1978. (ERIC Document Reproduction Service No. ED 161 000, 41p., PC-$3.65, MF-$.91)

Anderson, T. H., & Armbruster, B. B. Studying (No. 155), January 1980. (ERIC Document Reproduction Service No. ED 181 427, 48p., PC-$3.65, MF-$.91)

Anderson, T. H., Standiford, S. N., & Alessi, S. M. Computer Assisted Problem Solving in an Introductory Statistics Course (No. 56), August 1977. (ERIC Document Reproduction Service No. ED 146 563, 26p., PC-$3.65, MF-$.91)

Anderson, T. H., Wardrop, J. L., Hively, W., Muller, K. E., Anderson, R. I., Hastings, C. N., & Fredericksen, J. Development and Trial of a Model for Developing Domain Referenced Tests of Reading Comprehension (No. 86), May 1978. (ERIC Document Reproduction Service No. ED 157 036, 69p., PC-$5.30, MF-$.91)

Andre, M. E. D. A., & Anderson, T. H. The Development and Evaluation of a Self-Questioning Study Technique (No. 87), June 1978. (ERIC Document Reproduction Service No. ED 157 037, 37p., PC-$3.65, MF-$.91)

Antos, S. J. Processing Facilitation in a Lexical Decision Task (No. 113), January 1979. (ERIC Document Reproduction Service No. ED 165 129, 84p., PC-$6.95, MF-$.91)

Armbruster, B. B. Learning Principles from Prose: A Cognitive Approach Based on Schema Theory (No. 11), July 1976. (ERIC Document Reproduction Service No. ED 134 934, 48p., PC-$3.65, MF-$.91)

Armbruster, B. B., & Anderson, T. H. The Effect of Mapping on the Free Recall of Expository Text (No. 160), February 1980. (ERIC Document Reproduction Service No. ED 182 735, 49p., PC-$3.65, MF-$.91)

Armbruster, B. B., Stevens, R. J., & Rosenshine, B. Analyzing Content Coverage and Emphasis: A Study of Three Curricula and Two Tests (No. 26), March 1977. (ERIC Document Reproduction Service No. ED 136 238, 22p., PC-$2.00, MF-$.91)

Arter, J. A., & Jenkins, J. R. Differential Diagnosis-Prescriptive Teaching: A Critical Appraisal (No. 80), January 1978. (ERIC Document Reproduction Service No. ED 150 578, 104p., PC-$8.60, MF-$.91)

Asher, S. R. Referential Communication (No. 90), June 1978. (ERIC Document Reproduction Service No. ED 159 597, 71p., PC-$5.30, MF-$.91)

Asher, S. R. Influence of Topic Interest on Black Children and White Children's Reading Comprehension (No. 99), July 1978. (ERIC Document Reproduction Service No. ED 159 661, 35p., PC-$3.65, MF-$.91)

Asher, S. R., Hymel, S., & Wigfield, A. Children's Comprehension of High- and Low-Interest Material and a Comparison of Two Cloze Scoring Methods (No. 17), November 1976. (ERIC Document Reproduction Service No. ED 134 939, 32p., PC-$3.65, MF-$.91)

Asher, S. R., & Wigfield, A. Influence of Comparison Training on Children's Referential Communication (No. 139), August 1979. (ERIC Document Reproduction Service No. ED 177 493, 42p., PC-$3.65, MF-$.91)

Asher, S. R., & Wigfield, A. Training Referential Communication Skills (No. 175), July 1980. (ERIC Document Reproduction Service No. ED 191 014, 54p., PC-$5.30, MF-$.91)

Baker, L. Processing Temporal Relationships in Simple Stories: Effects of Input Sequence (No. 84), April 1978. (ERIC Document Reproduction Service No. ED 157 016, 54p., PC-$5.30, MF-$.91)

Baker, L. Comprehension Monitoring: Identifying and Coping with Text Confusions (No. 145), September 1979. (ERIC Document Reproduction Service No. ED 177 525, 62p., PC-$5.30, MF-$.91)

Baker, L., & Anderson, R. I. Effects of Inconsistent Information on Text Processing: Evidence for Comprehension Monitoring (No. 203), May 1981.

Baker, L., & Brown, A. L. Metacognitive Skills and Reading (No. 188), November 1980. (ERIC Document Reproduction Service No. ED 195 932, 74p., PC-$5.30, MF-$.91)

Baker, L., & Stein, N. L. The Development of Prose Comprehension Skills (No. 102), September 1978. (ERIC Document Reproduction Service No. ED 159 663, 69p., PC-$5.30, MF-$.91)

Barnitz, J. Interrelationship of Orthography and Phonological Structure in Learning to Read (No. 57), August 1977. (ERIC Document Reproduction Service No. ED 150 546, 62p., PC-$5.30, MF-$.91)

Barnitz, J. G. Reading Comprehension of Pronoun-Referent Structures by Children in Grades Two, Four, and Six (No. 117), March 1979. (ERIC Document Reproduction Service No. ED 170 731, 51p., PC-$5.30, MF-$.91)

Brewer, W. F. Memory for the Pragmatic Implications of Sentences (No. 65), October 1977. (ERIC Document Reproduction Service No. ED 146 564, 27p., PC-$3.65, MF-$.91)

Brewer, W. F., & Lichtenstein, E. H. Event Schemas, Story Schemas, and Story Grammars (No. 197), December 1980. (ERIC Document Reproduction Service No. ED 199 668, 46p., PC-$3.65, MF-$.91)

Brown, A. L. Knowing When, Where, and How to Remember: A Problem of Metacognition (No. 47), June 1977. (ERIC Document Reproduction Service No. ED 146 562, 152p., PC-$11.90, MF-$.91)

Brown, A. L. Theories of Memory and the Problems of Development: Activity, Growth, and Knowledge (No. 51), July 1977. (ERIC Document Reproduction Service No. ED 144 041, 59p., PC-$5.30, MF-$.91)

Brown, A. L. Learning and Development: The Problems of Compatibility, Access, and Induction (No. 165), March 1980. (ERIC Document Reproduction Service No. ED 184 093, 76p., PC-$6.95, MF-$.91)

Brown, A. L., & Campione, J. C. Memory Strategies in Learning: Training Children to Study Strategically (No. 22), March 1977. (ERIC Document Reproduction Service No. ED 136 234, 54p., PC-$5.30, MF-$.91)

Brown, A. L., & Campione, J. C. Permissible Inferences from the Outcome of Training Studies in Cognitive Development Research (No. 127), May 1979. (ERIC Document Reproduction Service No. ED 170 736, 34p., PC-$3.65, MF-$.91)

Brown, A. L., & Campione, J. C. Inducing Flexible Thinking: The Problem of Access (No. 156), January 1980. (ERIC Document Reproduction Service No. ED 181 428, 44p., PC-$3.65, MF-$.91)

Brown, A. L., Campione, J. C., & Barclay, C. R. Training Self-Checking Routines for Estimating Test Readiness: Generalization from List Learning to Prose Recall (No. 94), July 1978. (ERIC Document Reproduction Service No. ED 158 226, 41p., PC-$3.65, MF-$.91)

Brown, A. L., Campione, J. C., & Day, J. D. Learning to Learn: On Training Students to Learn from Texts (No. 189), November 1980.

Brown, A. L., & DeLoache, J. S. Skills, Plans, and Self-Regulation (No. 48), July 1977. (ERIC Document Reproduction Service No. ED 144 040, 66p., PC-$5.30, MF-$.91)

Brown, A. L., & French, L. A. The Zone of Potential Development: Implications for Intelligence Testing in the Year 2000 (No. 128), May 1979. (ERIC Document Reproduction Service No. ED 170 737, 46p., PC-$3.65, MF-$.91)

Brown, A. L., & Smiley, S. S. The Development of Strategies for Studying Prose Passages (No. 66), October 1977. (ERIC Document Reproduction Service No. ED 145 371, 59p., PC-$5.30, MF-$.91)

Brown, A. L., Smiley, S. S., Day, J. D., Townsend, M. A. R., & Lawton, S. C. Intrusion of a Thematic Idea in Children's Comprehension and Retention of Stories (No. 18), December 1976. (ERIC Document Reproduction Service No. ED 136 189, 39p., PC-$3.65, MF-$.91)

Brown, A. L., Smiley, S. S., & Lawton, S. C. The Effects of Experience on the Selection of Suitable Retrieval Cues for Studying from Prose Passages (No. 53), July 1977. (ERIC Document Reproduction Service No. ED 144 042, 30p., PC-$3.65, MF-$.91)

Bruce, B. C. Plans and Social Actions (No. 34), April 1977. (ERIC Document Reproduction Service No. ED 149 328, 45p., PC-$3.65, MF-$.91)

Bruce, B. Analysis of Interacting Plans as a Guide to the Understanding of Story Structure (No. 130), June 1979. (ERIC Document Reproduction Service No. ED 174 951, 43p., PC-$3.65, MF-$.91)

Bruce, B. C., Collins, A., Rubin, A. D., & Gentner, D. A Cognitive Science Approach to Writing (No. 89), June 1978. (ERIC Document Reproduction Service No. ED 157 039, 57p., PC-$5.30, MF-$.91)

Bruce, B. C., & Newman, D. Interacting Plans (No. 88), June 1978. (ERIC Document Reproduction Service No. ED 157 038, 100p., PC-$6.95, MF-$.91)

Campione, J. C., Nitsch, K., Bray, N., & Brown, A. L. Improving Memory Skills in Mentally Retarded Children: Empirical Research and Strategies for Intervention (No. 196), December 1980. (ERIC Document Reproduction Service No. ED 199 667, 67p., PC-$5.30, MF-$.91)

Canney, G., & Winograd, P. Schemata for Reading and Reading Comprehension Performance (No. 120), April 1979. (ERIC Document Reproduction Service No. ED 169 520, 99p., PC-$6.95, MF-$.91)

Cohen, P. R., & Perrault, C. R. Elements of a Plan-Based Theory of Speech Acts (No. 141), September 1979. (ERIC Document Reproduction Service No. ED 177 497, 76p., PC-$6.95, MF-$.91)

Collins, A., Brown, A. L., Morgan, J. L., & Brewer, W. F. The Analysis of Reading Tasks and Texts (No. 43), April 1977. (ERIC Document Reproduction Service No. ED 145 404, 96p., PC-$6.95, MF-$.91)

Collins, A., Brown, J. S., & Larkin, K. M. Inference in Text Understanding (No. 40), December 1977. (ERIC Document Reproduction Service No. ED 150 547, 48p., PC-$3.65, MF-$.91)

Collins, A., & Smith, E. E. Teaching the Process of Reading Comprehension (No. 182), September 1980. (ERIC Document Reproduction Service No. ED 193 616, 43p., PC-$3.65, MF-$.91)

Davison, A. Linguistics and the Measurement of Syntactic Complexity: The Case of Raising (No. 173), May 1980. (ERIC Document Reproduction Service No. ED 186 848, 60p., PC-$5.30, MF-$.91)

Davison, A., Kantor, R. N., Hannah, J., Hermon, G., Lutz, R., Salzillo, R. Limitations of Readability Formulas in Guiding Adaptations of Texts (No. 162), March 1980. (ERIC Document Reproduction Service No. ED 184 090, 157p., PC-$11.90, MF-$.91)

Dunn, B. R., Gould, J. E., & Singer, M. Cognitive Style Differences in Expository Prose Recall (No. 210), July 1981.

Dunn, B. R., Mathews, S. R., II, & Bieger, G. Individual Differences in the Recall of Lower-Level Textual Information (No. 150), December 1979. (ERIC Document Reproduction Service No. ED 181 448, 37p., PC-$3.65, MF-$.91)

Durkin, D. What Classroom Observations Reveal about Reading Comprehension Instruction (No. 106), October 1978. (ERIC Document Reproduction Service No. ED 162 259, 94p., PC-$6.95, MF-$.91)

Fleisher, L. S., & Jenkins, J. R. Effects of Contextualized and Decontextualized Practice Conditions on Word Recognition (No. 54), July 1977. (ERIC Document Reproduction Service No. ED 144 043, 37p., PC-$3.65, MF-$.91)

Fleisher, L. S., Jenkins, J. R., & Pany, D. Effects on Poor Readers' Comprehension of Training in Rapid Decoding (No. 103), September 1978. (ERIC Document Reproduction Service No. ED 159 664, 39p., PC-$3.65, MF-$.91)

Freebody, P., & Anderson, R. C. Effects of Differing Proportions and Locations of Difficult Vocabulary on Text Comprehension (No. 202), May 1981.

Gearhart, M., & Hall, W. S. Internal State Words: Cultural and Situational Variation in Vocabulary Usage (No. 115), February 1979. (ERIC Document Reproduction Service No. ED 165 131, 66p., PC-$5.30, MF-$.91)

Gentner, D. On Relational Meaning: The Acquisition of Verb Meaning (No. 78), December 1977. (ERIC Document Reproduction Service No. ED 149 325, 46p., PC-$3.65, MF-$.91)

Gentner, D. Semantic Integration at the Level of Verb Meaning (No. 114), February 1979. (ERIC Document Reproduction Service No. ED 165 130, 39p., PC-$3.65, MF-$.91)

Gentner, D. Verb Semantic Structures in Memory for Sentences: Evidence for Componential Representation (No. 151), December 1979. (ERIC Document Reproduction Service No. ED 181 424, 75p., PC-$5.30, MF-$.91)

Geva, E. Facilitating Reading Comprehension through Flowcharting (No. 211), July 1981.

Goetz, E. T. Sentences in Lists and in Connected Discourse (No. 3), November 1975. (ERIC Document Reproduction Service No. ED 134 927, 75p., PC-$5.30, MF-$.91)

Goetz, E. T. Inferences in the Comprehension of and Memory for Text (No. 49), July 1977. (ERIC Document Reproduction Service No. ED 150 548, 97p., PC-$6.95, MF-$.91)

Goetz, E. T., Anderson, R. C., & Schallert, D. L. The Representation of Sentences in Memory (No. 144), September 1979. (ERIC Document Reproduction Service No. ED 177 527, 71p., PC-$5.30, MF-$.91)

Goetz, E. T., & Osborn, J. Procedures for Sampling Texts and Tasks in Kindergarten through Eighth Grade (No. 30), April 1977. (ERIC Document Reproduction Service No. ED 146 565, 80p., PC-$6.95, MF-$.91)

Green, G. M. Discourse Functions of Inversion Construction (No. 98), July 1978. (ERIC Document Reproduction Service No. ED 160 998, 42p., PC-$3.65, MF-$.91)

Green, G. M. Organization, Goals, and Comprehensibility in Narratives: Newswriting, a Case Study (No. 132), July 1979. (ERIC Document Reproduction Service No. ED 174 949, 66p., PC-$5.30, MF-$.91)

Green, G. M. Linguistics and the Pragmatics of Language Use: What You Know When You Know a Language . . . and What Else You Know (No. 179), August 1980. (ERIC Document Reproduction Service No. ED 193 666, 73p., PC-$5.30, MF-$.91)

Green, G. M., Kantor, R. N., Morgan, J. L., Stein, N. L., Hermon, G., Salzillo, R., & Sellner, M. B. Analysis of "Babar Loses His Crown" (No. 169), April 1980. (ERIC Document Reproduction Service No. ED 185 514, 89p., PC-$6.95, MF-$.91)

Green, G. M., Kantor, R. N., Morgan, J. L., Stein, N. L., Hermon, G., Salzillo, R., & Sellner, M. B. Analysis of "The Wo[illegible]rful Desert" (No. 170), April 1980. (ERIC Document Reproduction Service No. ED 185 515, 47p., PC-$3.65, MF-$.91)

Green, G. M., Kantor, R. N., Morgan, J. L., Stein, N. L., Hermon, G., Salzillo, R., Sellner, M. B., Bruce, B. C., Gentner, D., & Webber, B. L. Problems and Techniques of Text Analysis (No. 168), April 1980. (ERIC Document Reproduction Service No. ED 185 513, 173p., PC-$11.90, MF-$.91)

Green, G. M., & Laff, M. O. Five-Year-Olds' Recognition of Authorship by Literary Style (No. 181), September 1980. (ERIC Document Reproduction Service No. ED 193 615, 44p., PC-$3.65, MF-$.91)

Grueneich, R., & Trabasso, T. The Story as Social Environment: Children's Comprehension and Evaluation of Intentions and Consequences (No. 142), September 1979. (ERIC Document Reproduction Service No. ED 177 496, 56p., PC-$5.30, MF-$.91)

Halff, H. M. Graphical Evaluation of Hierarchical Clustering Schemes (No. 1), October 1975. (ERIC Document Reproduction Service No. ED 134 926, 11p., PC-$2.00, MF-$.91)

Hall, W. S., & Dore, J. Lexical Sharing in Mother-Child Interaction (No. 161), March 1980. (ERIC Document Reproduction Service No. ED 184 066, 39p., PC-$3.65, MF-$.91)

Hall, W. S., & Guthrie, L. F. On the Dialect Question and Reading (No. 121), May 1979. (ERIC Document Reproduction Service No. ED 169 522, 32p., PC-$3.65, MF-$.91)

Hall, W. S., & Guthrie, L. F. Cultural and Situational Variation in Language Function and Use: Methods and Procedures for Research (No. 148), October 1979. (ERIC Document Reproduction Service No. ED 179 944, 49p., PC-$3.65, MF-$.91)

Hall, W. S., Linn, R. L., & Nagy, W. E. Spoken Words (No. 177), August 1980.

Hall, W. S., & Nagy, W. E. Theoretical Issues in the Investigation of Words of Internal Report (No. 146), October 1979. (ERIC Document Reproduction Service No. ED 177 526, 108p., PC-$8.60, MF-$.91)

Hall, W. S., Nagy, W. E., & Nottenburg, G. Situational Variation in the Use of Internal State Words (No. 212), August 1981.

Hall, W. S., & Tirre, W. C. The Communicative Environment of Young Children: Social Class, Ethnic, and Situational Differences (No. 125), May 1979. (ERIC Document Reproduction Service No. ED 170 788, 30p., PC-$3.65, MF-$.91)

Hansen, J., & Pearson, P. D. The Effects of Inference Training and Practice on Young Children's Comprehension (No. 166), April 1980. (ERIC Document Reproduction Service No. ED 186 839, 53p., PC-$5.30, MF-$.91)

Hayes, D. A., & Tierney, R. J. Increasing Background Knowledge through Analogy: Its Effects upon Comprehension and Learning (No. 186), October 1980. (ERIC Document Reproduction Service No. ED 195 953, 81p., PC-$6.95, MF-$.91)

Hermon, G. On the Discourse Structure of Direct Quotation (No. 143), September 1979. (ERIC Document Reproduction Service No. ED 177 495, 46p., PC-$3.65, MF-$.91)

Hogaboam, T. W., & McConkie, G. W. The Rocky Road from Eye Fixations to Comprehension (No. 207), May 1981.

Huggins, A. W. F. Syntactic Aspects of Reading Comprehension (No. 33), April 1977. (ERIC Document Reproduction Service No. ED 142 972, 68p., PC-$5.30, MF-$.91)

Iran-Nejad, A. The Schema: A Structural or a Functional Pattern (No. 159), February 1980. (ERIC Document Reproduction Service No. ED 181 449, 46p., PC-$3.65, MF-$.91)

Iran-Nejad, A., Ortony, A., & Rittenhouse, R. K. The Comprehension of Metaphorical Uses of English by Deaf Children (No. 184), October 1980. (ERIC Document Reproduction Service No. ED 193 618, 34p., PC-$3.65, MF-$.91)

Jenkins, J. R., & Larson, K. Evaluating Error Correction Procedures for Oral Reading (No. 55), June 1978. (ERIC Document Reproduction Service No. ED 158 224, 34p., PC-$3.65, MF-$.91)

Jenkins, J. R., & Pany, D. Curriculum Biases in Reading Achievement Tests (No. 16), November 1976. (ERIC Document Reproduction Service No. ED 134 938, 24p., PC-$2.00, MF-$.91)

Jenkins, J. R., Pany, D., & Schreck, J. Vocabulary and Reading Comprehension: Instructional Effects (No. 100), August 1978. (ERIC Document Reproduction Service No. ED 160 999, 50p., PC-$3.65, MF-$.91)

Johnston, P. Implications of Basic Research for the Assessment of Reading Comprehension (No. 206), May 1981.

Kane, J. H., & Anderson, R. C. Depth of Processing and Interference Effects in the Learning and Remembering of Sentences (No. 21), February 1977. (ERIC Document Reproduction Service No. ED 134 942, 29p., PC-$3.65, MF-$.91)

Kleiman, G. M. The Effect of Previous Context on Reading Individual Words (No. 20), February 1977. (ERIC Document Reproduction Service No. ED 134 941, 76p., PC-$6.95, MF-$.91)

Kleiman, G. M. The Prelinguistic Cognitive Basis of Children's Communicative Intentions (No. 19), February 1977. (ERIC Document Reproduction Service No. ED 134 940, 51p., PC-$5.30, MF-$.91)

Kleiman, G. M. The Scope of Facilitation of Word Recognition from Single Word and Sentence Frame Contexts (No. 133), July 1979. (ERIC Document Reproduction Service No. ED 174 947, 61p., PC-$5.30, MF-$.91)

Kleiman, G. M., Winograd, P. N., & Humphrey, M. M. Prosody and Children's Parsing of Sentences (No. 123), May 1979. (ERIC Document Reproduction Service No. ED 170 733, 28p., PC-$3.65, MF-$.91)

Linn, R. L., Levine, M. V., Hastings, C. N., & Wardrop, J. L. An Investigation of Item Bias in a Test of Reading Comprehension (No. 163), March 1980. (ERIC Document Reproduction Service No. ED 184 091, 97p., PC-$6.95, MF-$.91)

Mason, J. M. Questioning the Notion of Independent Processing Stages in Reading (No. 8), February 1976. (Journal of Educational Psychology, 1977, 69, 288-297.

Mason, J. M. Reading Readiness: A Definition and Skills Hierarchy from Preschoolers' Developing Conceptions of Print (No. 59), September 1977. (ERIC Document Reproduction Service No. ED 145 403, 57p., PC-$5.30, MF-$.91)

Mason, J. M. The Role of Strategy in Reading in the Mentally Retarded (No. 58), September 1977. (ERIC Document Reproduction Service No. ED 145 406, 28p., PC-$3.65, MF-$.91)

Mason, J. M. Prereading: A Developmental Perspective (No. 198), February 1981. (ERIC Document Reproduction Service No. ED 199 659, 96p., PC-$6.95, MF-$.91)

Mason, J. M., & Au, K. H. Learning Social Context Characteristics in Prereading Lessons (No. 205), May 1981.

Mason, J. M., & Kendall, J. R. Facilitating Reading Comprehension Through Text Structure Manipulation (No. 92), June 1978. (ERIC Document Reproduction Service No. ED 157 041, 36p., PC-$3.65, MF-$.91)

Mason, J. M., Knisely, E., & Kendall, J. Effects of Polysemous Words on Sentence Comprehension (No. 85), May 1978. (ERIC Document Reproduction Service No. ED 157 015, 34p., PC-$3.65, MF-$.91)

Mason, J., & McCormick, C. Testing the Development of Reading and Linguistic Awareness (No. 126), May 1979. (ERIC Document Reproduction Service No. ED 170 735, 50p., PC-$3.65, MF-$.91)

Mason, J., Osborn, J., & Rosenshine, B. A Consideration of Skill Hierarchy Approaches to the Teaching of Reading (No. 42), December 1977. (ERIC Document Reproduction Service No. ED 150 549, 176p., PC-$13.55, MF-$.91)

McClure, E. Aspects of Code-Switching in the Discourse of Bilingual Mexican-American Children (No. 44), April 1977. (ERIC Document Reproduction Service No. ED 142 975, 38p., PC-$3.65, MF-$.91)

McClure, E., Mason, J., & Barnitz, J. Story Structure and Age Effects on Children's Ability to Sequence Stories (No. 122), May 1979. (ERIC Document Reproduction Service No. ED 170 732, 75p., PC-$5.30, MF-$.91)

McClure, E., Mason, J., & Williams, J. Sociocultural Variables in Children's Sequencing of Stories (No. 209), July 1981.

McClure, E., & Steffensen, M. S. A Study of the Use of Conjunctions across Grades and Ethnic Groups (No. 158), January 1980. (ERIC Document Reproduction Service No. ED 182 688, 43p., PC-$3.65, MF-$.91)

McConkie, G. W. Evaluating and Reporting Data Quality in Eye Movement Research (No. 193), December 1980. (ERIC Document Reproduction Service No. ED 199 664, 50p., PC-$3.65, MF-$.91)

McConkie, G. W., Hogaboam, T. W., Wolverton, G. S., Zola, D., & Lucas, P. A. Toward the Use of Eye Movements in the Study of Language Processing (No. 134), August 1979. (ERIC Document Reproduction Service No. ED 174 968, 48p., PC-$3.65, MF-$.91)

McConkie, G. W., & Zola, D. Language Constraints and the Functional Stimulus in Reading (No. 194), December 1980. (ERIC Document Reproduction Service No. ED 199 665, 51p., PC-$5.30, MF-$.91)

Morgan, J. L. Two Types of Convention in Indirect Speech Acts (No. 52), July 1977. (ERIC Document Reproduction Service No. ED 145 405, 40p., PC-$3.65, MF-$.91)

Nash-Webber, B. Anaphora: A Cross-Disciplinary Survey (No. 31), April 1977. (ERIC Document Reproduction Service No. ED 144 039, 43p., PC-$3.65, MF-$.91)

Nash-Webber, B. L. Inferences in an Approach to Discourse Anaphora (No. 77), January 1978. (ERIC Document Reproduction Service No. ED 150 552, 30p., PC-$3.65, MF-$.91)

Nash-Webber, B., & Reiter, R. Anaphora and Logical Form: On Formal Meaning Representation for Natural Language (No. 36), April 1977. (ERIC Document Reproduction Service No. ED 142 973, 42p., PC-$3.65, MF-$.91)

Nezworski, T., Stein, N. L., & Trabasso, T. Story Structure Versus Content Effects on Children's Recall and Evaluative Inferences (No. 129), June 1979. (ERIC Document Reproduction Service No. ED 172 187, 49p., PC-$3.65, MF-$.91)

Nicholson, T., Pearson, P. D., & Dykstra, R. Effects of Embedded Anomalies and Oral Reading Errors on Children's Understanding of Stories (No. 118), March 1979. (ERIC Document Reproduction Service No. ED 169 524, 43p., PC-$3.65, MF-$.91)

Nolan, S. D., Tanenhaus, M. K., & Seidenberg, M. S. Multiple Code Activation in Word Recognition: Evidence from Rhyme Monitoring (No. 204), May 1981.

Ortony, A. Names, Descriptions, and Pragmatics (No. 7), February 1976. (ERIC Document Reproduction Service No. ED 134 931, 25p., PC-$2.00, MF-$.91)

Ortony, A. Remembering and Understanding Jabberwocky and Small-Talk (No. 28), March 1977. (ERIC Document Reproduction Service No. ED 137 753, 36p., PC-$3.65, MF-$.91)

Ortony, A. Beyond Literal Similarity (No. 105), October 1978. (ERIC Document Reproduction Service No. ED 166 635, 58p., PC-$5.30, MF-$.91)

Ortony, A. Some Psycholinguistic Aspects of Metaphor (No. 112), January 1979. (ERIC Document Reproduction Service No. ED 165 115, 38p., PC-$3.65, MF-$.91)

Ortony, A. Understanding Metaphors (No. 154), January 1980. (ERIC Document Reproduction Service No. ED 181 426, 52p., PC-$5.30, MF-$.91)

Ortony, A., Reynolds, R. E., & Arter, J. A. Metaphor: Theoretical and Empirical Research (No. 27), March 1977. (ERIC Document Reproduction Service No. ED 137 752, 63p., PC-$5.30, MF-$.91)

Ortony, A., Schallert, D. L., Reynolds, R. E., & Antos, S. J. Interpreting Metaphors and Idioms: Some Effects of Context on Comprehension (No. 93), July 1978. (ERIC Document Reproduction Service No. ED 157 042, 41p., PC-$3.65, MF-$.91)

Pany, D., & Jenkins, J. R. Learning Word Meanings: A Comparison of Instructional Procedures and Effects on Measures of Reading Comprehension with Learning Disabled Students (No. 25), March 1977. (ERIC Document Reproduction Service No. ED 136 237, 34p., PC-$3.65, MF-$.91)

Pearson, P. D., Hansen, J., & Gordon, C. The Effect of Background Knowledge on Young Children's Comprehension of Explicit and Implicit Information (No. 116), March 1979. (ERIC Document Reproduction Service No. ED 169 521, 26p., PC-$3.65, MF-$.91)

Pearson, P. D., Raphael, T., TePaske, N., & Hyser, C. The Function of Metaphor in Children's Recall of Expository Passages (No. 131), July 1979. (ERIC Document Reproduction Service No. ED 174 950, 41p., PC-$3.65, MF-$.91)

Pichert, J. W. Sensitivity to What is Important in Prose (No. 149), November 1979. (ERIC Document Reproduction Service No. ED 179 946, 64p., PC-$5.30, MF-$.91)

Pichert, J. W., & Anderson, R. C. Taking Different Perspectives on a Story (No. 14), November 1976. (ERIC Document Reproduction Service No. ED 134 936, 30p., PC-$3.65, MF-$.91)

Raphael, T. E., Myers, A. C., Freebody, P., Tirre, W. C., & Fritz, M. Contrasting the Effects of Some Text Variables on Comprehension and Ratings of Comprehensibility (No. 190), December 1980. (ERIC Document Reproduction Service No. ED 199 661, 58p., PC-$5.30, MF-$.91)

Reder, L. M. Comprehension and Retention of Prose: A Literature Review (No. 108), November 1978. (ERIC Document Reproduction Service No. ED 165 114, 116p., PC-$8.60, MF-$.91)

Reichman, R. Conversational Coherency (No. 95), July 1978. (ERIC Document Reproduction Service No. ED 159 658, 86p., PC-$6.95, MF-$.91)

Reynolds, R. E., & Anderson, R. C. Influence of Questions on the Allocation of Attention during Reading (No. 183), October 1980. (ERIC Document Reproduction Service No. ED 193 617, 44p., PC-$3.65, MF-$.91)

Reynolds, R. E., & Ortony, A. Some Issues in the Measurement of Children's Comprehension of Metaphorical Language (No. 172), May 1980. (ERIC Document Reproduction Service No. ED 185 542, 42p., PC-$3.65, MF-$.91)

Reynolds, R. E., Standiford, S. N., & Anderson, R. C. Distribution of Reading Time When Questions are Asked about a Restricted Category of Text Information (No. 83), April 1978. (ERIC Document Reproduction Service No. ED 153 206, 34p., PC-$3.65, MF-$.91)

Reynolds, R. E., Taylor, M. A., Steffensen, M. S., Shirey, L. L., & Anderson, R. C. Cultural Schemata and Reading Comprehension (No. 201), April 1981.

Royer, J. M. Theories of Learning Transfer (No. 79), January 1978. (ERIC Document Reproduction Service No. ED 149 326, 55p., PC-$5.30, MF-$.91)

Royer, J. M., & Cunningham, D. J. On the Theory and Measurement of Reading Comprehension (No. 91), June 1978. (ERIC Document Reproduction Service No. ED 157 040, 63p., PC-$5.30, MF-$.91)

Royer, J. M., Hastings, C. N., & Hook, C. A Sentence Verification Technique for Measuring Reading Comprehension (No. 137), August 1979. (ERIC Document Reproduction Service No. ED 176 234, 34p., PC-$3.65, MF-$.91)

Rubin, A. D. A Theoretical Taxonomy of the Differences between Oral and Written Language (No. 35), January 1978. (ERIC Document Reproduction Service No. ED 150 550, 61p., PC-$5.30, MF-$.91)

Rubin, A. D., Bruce, B. C., & Brown, J. S. A Process-Oriented Language for Describing Aspects of Reading Comprehension (No. 13), November 1976. (ERIC Document Reproduction Service No. ED 136 188, 41p., PC-$3.65, MF-$.91)

Schallert, D. L. Improving Memory for Prose: The Relationship between Depth of Processing and Context (No. 5), November 1975. (ERIC Document Reproduction Service No. ED 134 929, 37p., PC-$3.65, MF-$.91)

Schallert, D. L., Kleiman, G. M., & Rubin, A. D. Analyses of Differences between Written and Oral Language (No. 29), April 1977. (ERIC Document Reproduction Service No. ED 144 038, 33p., PC-$3.65, MF-$.91)

Schwartz, R. M. Strategic Processes in Beginning Reading (No. 15), November 1976. (ERIC Document Reproduction Service No. ED 134 937, 19p., PC-$2.00, MF-$.91)

Schwartz, R. M. Relation of Context Utilization and Orthographic Automaticity in Word Identification (No. 45), May 1977. (ERIC Document Reproduction Service No. ED 137 762, 27p., PC-$3.65, MF-$.91)

Schwartz, R. M. Levels of Processing: The Strategic Demands of Reading Comprehension (No. 135), August 1979. (ERIC Document Reproduction Service No. ED 177 471, 45p., PC-$3.65, MF-$.91)

Seidenberg, M. S., Tanenhaus, M. K., & Leiman, J. M. The Time Course of Lexical Ambiguity Resolution in Context (No. 164), March 1980. (ERIC Document Reproduction Service No. ED 184 092, 58p., PC-$5.30, MF-$.91)

Shimron, J., & Navon, D. The Dependence on Graphemes and on Their Translation to Phonemes in Reading: A Developmental Perspective (No. 208), June 1981.

Shoben, E. J. Choosing a Model of Sentence Picture Comparisons: A Reply to Catlin and Jones (No. 81), February 1978. (ERIC Document Reproduction Service No. ED 150 577, 30p., PC-$3.65, MF-$.91)

Shoben, E. J., Rips, L. J., & Smith, E. E. Issues in Semantic Memory: A Response to Glass and Holyoak (No. 101), August 1978. (ERIC Document Reproduction Service No. ED 159 662, 85p., PC-$6.95, MF-$.91)

Siegel, M. A. Teacher Behaviors and Curriculum Packages: Implications for Research and Teacher Education (No. 9), April 1976. (ERIC Document Reproduction Service No. ED 134 932, 42p., PC-$3.65, MF-$.91)

Smiley, S. S., Oakley, D. D., Worthen, D., Campione, J. C., & Brown, A. L. Recall of Thematically Relevant Material by Adolescent Good and Poor Readers as a Function of Written Versus Oral Presentation (No. 23), March 1977. (ERIC Document Reproduction Service No. ED 136 235, 23p., PC-$2.00, MF-$.91)

Smith, E. E. Organization of Factual Knowledge (No. 185), October 1980. (ERIC Document Reproduction Service No. ED 195 954, 109p., PC-$8.60, MF-$.91)

Spiro, R. J. Inferential Reconstruction in Memory for Connected Discourse (No. 2), October 1975. (ERIC Document Reproduction Service No. ED 136 187, 81p., PC-$6.95, MF-$.91)

Spiro, R. J. Etiology of Reading Comprehension Style (No. 124), May 1979. (ERIC Document Reproduction Service No. ED 170 734, 21p., PC-$2.00, MF-$.91)

Spiro, R. J. Prior Knowledge and Story Processing: Integration, Selection, and Variation (No. 138), August 1979. (ERIC Document Reproduction Service No. ED 176 235, 41p., PC-3.32, MF-$.91)

Spiro, R. J. Schema Theory and Reading Comprehension: New Directions (No. 191), December 1980. (ERIC Document Reproduction Service No. ED 199 662, 49p., PC-$3.65, MF-$.91)

Spiro, R. J., & Esposito, J. J. Superficial Processing of Explicit Inferences in Text (No. 60), December 1977. (ERIC Document Reproduction Service No. ED 150 545, 27p., PC-$3.65, MF-$.91)

Spiro, R. J., & Taylor, B. M. On Investigating Children's Transition from Narrative to Expository Discourse: The Multidimensional Nature of Psychological Text Classification (No. 195), December 1980. (ERIC Document Reproduction Service No. ED 199 666, 43p., PC-$3.65, MF-$.91)

Spiro, R. J., & Tirre, W. C. Individual Differences in Schema Utilization During Discourse Processing (No. 111), January 1979. (ERIC Document Reproduction Service No. ED 166 651, 29p., PC-$3.65, MF-$.91)

Steffensen, M. S. Bereiter and Engelmann Reconsidered: The Evidence from Children Acquiring Black English Vernacular (No. 82), March 1978. (ERIC Document Reproduction Service No. ED 153 204, 31p., PC-$3.65, MF-$.91)

Steffensen, M. S., & Guthrie, L. F. Effect of Situation on the Verbalization of Black Inner-City Children (No. 180), September 1980. (ERIC Document Reproduction Service No. ED 193 614, 37p., PC-$3.65, MF-$.91)

Steffensen, M. S., Jogdeo, C., & Anderson, R. C. A Cross-Cultural Perspective on Reading Comprehension (No. 97), July 1978. (ERIC Document Reproduction Service No. ED 159 660, 41p., PC-$3.65, MF-$.91)

Steffensen, M. S., Reynolds, R. E., McClure, E., & Guthrie, L. F. Black English Vernacular and Reading Comprehension: A Cloze Study of Third, Sixth, and Ninth Graders (No. 199), February 1981. (ERIC Document Reproduction Service No. ED 199 660, 44p., PC-$3.65, MF-$.91)

Stein, N. L. How Children Understand Stories: A Developmental Analysis (No. 69), March 1978. (ERIC Document Reproduction Service No. ED 153 205, 68p., PC-$5.30, MF-$.91)

Stein, N. L., & Goldman, S. Children's Knowledge about Social Situations: From Causes to Consequences (No. 147), October 1979. (ERIC Document Reproduction Service No. ED 177 524, 54p., PC-$5.30, MF-$.91)

Stein, N. L., & Nezworski, T. The Effects of Organization and Instructional Set on Story Memory (No. 68), January 1978. (ERIC Document Reproduction Service No. ED 149 327, 41p., PC-$3.65, MF-$.91)

Stein, N. L., & Trabasso, T. What's in a Story: An Approach to Comprehension and Instruction (No. 200), April 1981.

Straker, D. Y. Situational Variables in Language Use (No. 167), April 1980. (ERIC Document Reproduction Service No. ED 185 619, 49p., PC-$3.65, MF-$.91)

Tanenhaus, M. K., Flanigan, H., & Seidenberg, M. S. Orthographic and Phonological Activation in Auditory and Visual Word Recognition (No. 178), August 1980. (ERIC Document Reproduction Service No. ED 193 620, 46p., PC-$3.65, MF-$.91)

Tanenhaus, M. K., & Seidenberg, M. S. Discourse Context and Sentence Perception (No. 176), July 1980. (ERIC Document Reproduction Service No. ED 191 015, 45p., PC-$3.65, MF-$.91)

Thieman, T. J., & Brown, A. L. The Effects of Semantic and Formal Similarity on Recognition Memory for Sentences in Children (No. 76), November 1977. (ERIC Document Reproduction Service No. ED 150 551, 26p., PC-$3.65, MF-$.91)

Tierney, R. J., & Cunningham, J. W. Research on Teaching Reading Comprehension (No. 187), November 1980. (ERIC Document Reproduction Service No. ED 195 946, 125p., PC-$8.60, MF-$.91)

Tierney, R. J., & Mosenthal, J. Discourse Comprehension and Production: Analyzing Text Structure and Cohesion (No. 152), January 1980. (ERIC Document Reproduction Service No. ED 179 945, 84p., PC-$6.95, MF-$.91)

Tirre, W. C., Freebody, P., & Kaufman, K. Achievement Outcomes of Two Reading Programs: An Instance of Aptitude-Treatment Interaction (No. 174), June 1980. (ERIC Document Reproduction Service No. ED 193 619, 34p., PC-$3.65, MF-$.91)

Tirre, W. C., Manelis, L., & Leicht, K. L. The Effects of Imaginal and Verbal Strategies on Prose Comprehension in Adults (No. 110), December 1978. (ERIC Document Reproduction Service No. ED 165 116, 27p., PC-$3.65, MF-$.91)

Trabasso, T. On the Making of Inferences During Reading and Their Assessment (No. 157), January 1980. (ERIC Document Reproduction Service No. ED 181 429, 38p., PC-$3.65, MF-$.91)

Wardrop, J. L., Anderson, T. H., Hively, W., Anderson, R. I., Hastings, C. N., & Muller, K. E. A Framework for Analyzing Reading Test Characteristics (No. 109), December 1978. (ERIC Document Reproduction Service No. ED 165 117, 65p., PC-$5.30, MF-$.91)

Wigfield, A., & Asher, S. R. Age Differences in Children's Referential Communication Performance: An Investigation of Task Effects (No. 96), July 1978. (ERIC Document Reproduction Service No. ED 159 659, 31p., PC-$3.65, MF-$.91)

Winograd, P., & Johnston, P. Comprehension Monitoring and the Error Detection Paradigm (No. 153), January 1980. (ERIC Document Reproduction Service No. ED 181 425, 57p., PC-$5.30, MF-$.91)

Woods, W. A. Multiple Theory Formation in High-Level Perception (No. 38), April 1977. (ERIC Document Reproduction Service No. ED 144 020, 58p., PC-$5.30, MF-$.91)

Zehler, A. M., & Brewer, W. F. Acquisition of the Article System in English (No. 171), May 1980. (ERIC Document Reproduction Service No. ED 186 907, 51p., PC-$5.30, MF-$.91)